2016

TOWARDS A GENERIC ONTOLOGY FOR SOLAR IRRADIANCE FORECASTING

Abhilash Kantamneni
Michigan Technological University, akantamn@mtu.edu

TOWARDS A GENERIC ONTOLOGY FOR SOLAR IRRADIANCE FORECASTING

By

Abhilash Kantamneni

A REPORT

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Computer Science

MICHIGAN TECHNOLOGICAL UNIVERSITY

2016

This report has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Computer Science.

Department of Computer Science

Report Advisor: *Dr. Laura E. Brown*

Committee Member: *Dr. Steven Goldsmith*

Committee Member: *Mr. Jay Meldrum*

Department Chair: *Dr. Min Song*

Dedication

To the spirit of the American Dream

"Family a champion, advisor a champion,

Colleagues a champion, friends a champion;

Everybody know funding agency a champion.

Don't vex if your name not call."

- Set to the tune of *Champion*, anthem of the 2016 Cricket World-Cup winning West Indies team.

Contents

List of Figures

List of Tables

Abstract

The growth of solar energy resources in recent years has led to increased calls for accurate forecasts of solar irradiance for the reliable and sustainable integration of solar into the national grid. A growing body of academic research has developed models for forecasting solar irradiance, identified metrics for comparing solar forecasts, and described applications and end users of solar forecasts.

In recent years, many disciplines are developing ontologies to facilitate better communication, improve inter-operabiity and refine knowledge reuse by experts and users of the domain. Ontologies are explicit and formal vocabulary of terms and their relationships. This report describes a step towards using ontologies to describe the knowledge, concepts and relationships in the domain of solar irradiance forecasting to develop a shared understanding for diverse stakeholders that interact with the domain. A preliminary ontology on solar irradiance forecasting was created and validated on three use cases.

Chapter 1

Introduction

Spurred by declining photovoltaic (PV) module prices, favorable government policies, and growing concerns about mitigating climate change, recent years have seen a rapid growth in the proliferation of solar electric systems. Since 2006, the installed cost of solar PV systems have dropped by 73% while in the same time period, the total installed capacity of solar in the US has increased by a staggering 9,900% [1] (see Fig. 1.1). In 2015, for the first time in U.S. history, more solar PV was added to the electric grid than natural gas fired generation. By 2020, solar energy is expected to be cost-competitive with other forms of electricity, even without subsidies [13]. Energy from solar technologies are expected to provide nearly a quarter of the world's electricity, by the year 2050 [14].

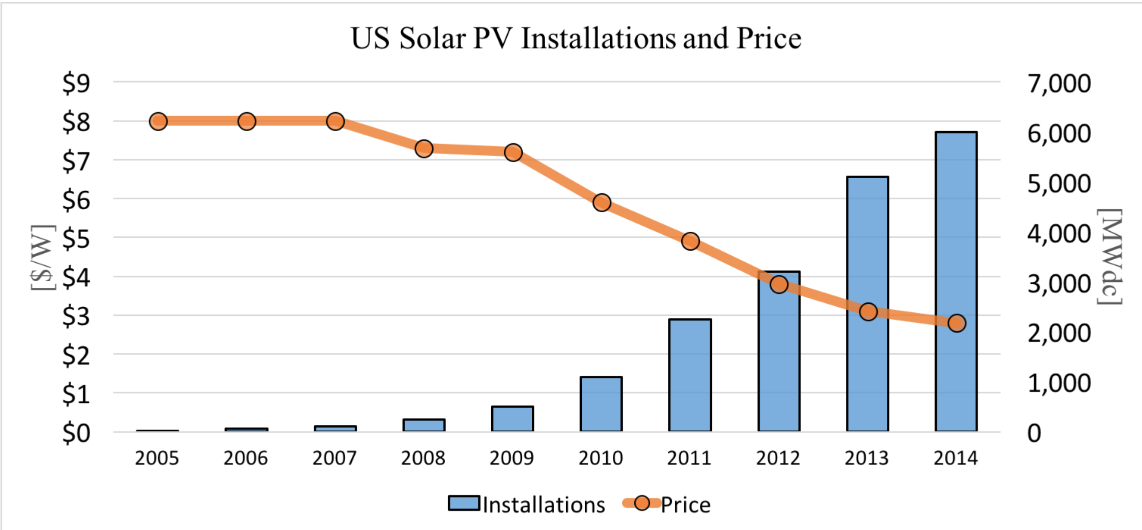

Figure 1.1: For the last decade, falling installed cost of solar ($ per kW) have coincided with increase in solar deployed (MW installed) on the grid. Adapted from [1]

As PV markets continue to grow, and the installed cost as well as the total levelized cost of energy (LCOE) from solar continues to decrease, the reliable and sustainable integration of solar into the national grid becomes a challenging problem. Electricity is unique as a commodity - through a combination of generation, transmission and distribution, electricity has to be made available the instant it needs to be consumed. Uncertainty in consumer demand, generation system outages, and transmission congestion create everyday challenges for grid operators and utilities. The *variable*, *intermittent* and *non-dispatchable* nature of solar energy introduces additional uncertainty and variability in grid operations [11].

At current low levels of solar energy generation connected to the grid, solar variability can be mitigated by using ancillary generator backups. For the electric grid of the future with significantly high penetration of solar electric generation, such methods may not be reliable, affordable or sustainable [15].

To address this, studies and industrial reports have called for high-precision solar power forecasting that can provide value to participants in the electric grid value chain [7, 8, 10, 12, 16]. Short-term (minutes-few hours) solar power forecasts are essential for power plant operations, grid balancing, real-time unit dispatching, trading on energy markets. Medium-term (few hours-day ahead) aid in unit commitment, reducing idle backup capacity, demand scheduling and reducing transmission congestion. Longer range forecasts are useful for resource adequacy planning and energy policy objectives [17].

The power and energy produced by solar electric systems depends on the total amount of solar irradiance incident at the location of the installation. Solar irradiance, in turn, depends on time, location, meteorological and atmospheric conditions. Most research is focused on forecasting solar irradiance, while using a mathematical model that accounts for the electrical, material, and orientation characteristics of the installation to calculate solar power forecast (see Fig 1.2).

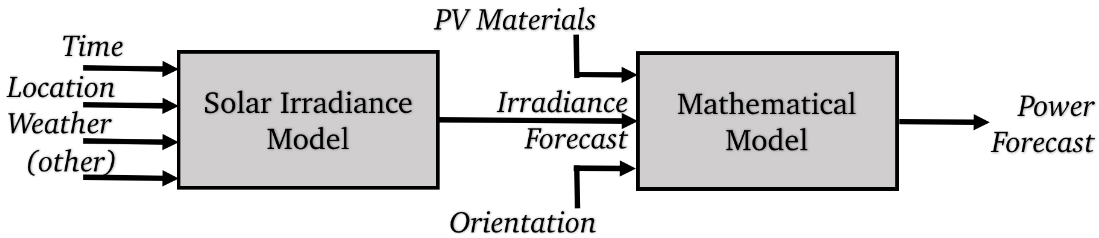

Figure 1.2: Solar power output from an installation can be generated by means of a solar irradiance forecast, physical characteristics of the installation and a simple mathematical model. Adapted in part from [2]

1.1 Solar Irradiance & Forecasting

Research in solar irradiance forecasting aggregates diverse areas of knowledge including atmospheric physics, remote sensing, forecasting theory and machine learning. Solar forecast models range in complexity - from simple location-specific geometrical formulations to complex ensemble assemblages of sky cameras, atmospheric sensors, and satellite imagery, each with its own dimensions of temporal and spatial applicability (see Table 2.2).

Models can also be classified based on their input characteristics and output specifications, instrumentation requirements, data availability and resolution. These attributes are neither distinct, nor exclusive. Indeed, models may share common attributes. Models also have specific temporal and spatial domains.

Models can also be evaluated on the basis of their performance - accuracy, specificity, precision, responsiveness and latency. Industry standards for forecast metrics and validation are slowly emerging [2, 4], with broad goals of producing more reliable forecasts, with a realistic expectation of forecast precision.

Solar irradiance itself can be resolved into individual components based on the path of incoming radiation through the atmosphere. Individual components of solar irradiance may have diverse and independent applications, each with different end users.

The growing diversity in forecast models, inputs, outputs, performance characteristics, instrumentation requirements, applications, temporal scales, spatial scales, and end users requires careful organization and representation of knowledge about solar irradiance forecasting.

1.2 Ontologies

In recent years, ontologies have emerged as a way to represent knowledge of a particular domain. Ontology - a term borrowed from philosophy, presently has wide applications in computer science, artificial intelligence and knowledge representation communities. Ontologies are "an explicit and formal specification of a conceptualization" [18], representing a set of concepts, events and relations that are specified to create a vocabulary for a domain. Computational ontologies can formally model a system, its constituent entities and relationships among them [19].

Modern semantic ontologies can facilitate sharing common understanding of structure of information between communities of interest, either human or software agents. Ontologies also allow reuse of domain knowledge. Large and complex ontologies can be built by integrating existing and well-defined ontologies. By separating domain knowledge from operational, ontologies promote inter-operability, translating between different methods, models and paradigms [20].

This report is a step towards representing implicit and explicit domain knowledge of solar radiation modeling and forecasting using ontologies in anticipation of a growing market in solar energy generation. A thorough literature review reveals no comprehensive semantic ontology to represent information and knowledge about solar irradiance forecasting.

A formal representation and informatic systems can reduce data uncertainty and improve the model selection process as a function of the constraints imposed by different operational conditions [21] . The continual modernization of the electric power grid through integration of digital and information technologies with dynamic distributed energy resources underscores the need for a formal ontology for solar forecast modeling [22].

The thesis will briefly survey the categories of models for solar forecasting: Clear

Sky, Parameterized, Numerical Weather Models, Stochastic Models, ANN models and Persistence Models) and review ontological model development methodologies: METHODONTOLOGY [23], Ontology 101 [20] , SENSUS[24] . A formal ontology for solar forecasting will be developed using free and open-source ontology editor, Prot*ég*é software [25]. The model will be evaluated by testing against select models from solar forecast meta-surveys.

Chapter 2

Solar Forecasting

This chapter introduces the basics of solar irradiance, and reviews widely used models for forecasting solar irradiance. Subsequently, metrics for evaluating solar forecast, applications and intended end-users of solar forecasting are briefly reviewed.

2.1 Solar Irradiance Basics

Irradiance is solar, short-wave radiation flux incident on Earth[26]. The total short-wave radiation received by a horizontal collector on the surface of the Earth is a sum of many parts - beam, diffuse and albedo. (see Fig. 2.1)

Beam irradiance is the irradiance that is transmitted directly from the sun to the incident horizontal surface in a straight-line path. Beam irradiance is sometimes referred to as Direct Normal Irradiance (DNI). Diffuse irradiance is the sum of all other scattered solar radiation that falls onto the horizontal incident surface. The diffuse component consists of radiation scattered off of the atmospheric molecules, particles and clouds. Diffuse irradiance is synonymous with Diffuse Horizontal Irradiance (DHI) [26, 27]. Albedo is the irradiance that is reflected by the ground and objects on the ground. Global Horizontal Irradiance (GHI) is the sum of all irradiance. Albedo effects are insignificant compared to beam and diffuse irradiance, therefore GHI is essentially the sum of DNI and DHI.

Extraterrestrial Horizontal Irradiance (EHI) is the irradiance measured just outside the Earths atmosphere on a plane tangential to the atmosphere, as shown in Figure 2.1. Extraterrestrial Horizontal Irradiance can be accurately calculated to a high degree of precision, as it is primarily a function of solar distance and global position [26] [28].

The clear sky irradiance is the maximum GHI incident on Earth, measured during periods of no cloud cover. Many models do not directly utilize the clear sky irradiance, but rather make use of the clear sky index, which is the ratio of GHI during overcast conditions to the clear sky irradiance [29].

Concentrated Solar Power (CSP) systems and dual-axis trackers, i.e. PV panels that follow the position of the sun through the day, are sensitive to DNI. Most PV panels for residential or small scale applications are sensitive to both GHI and DNI.

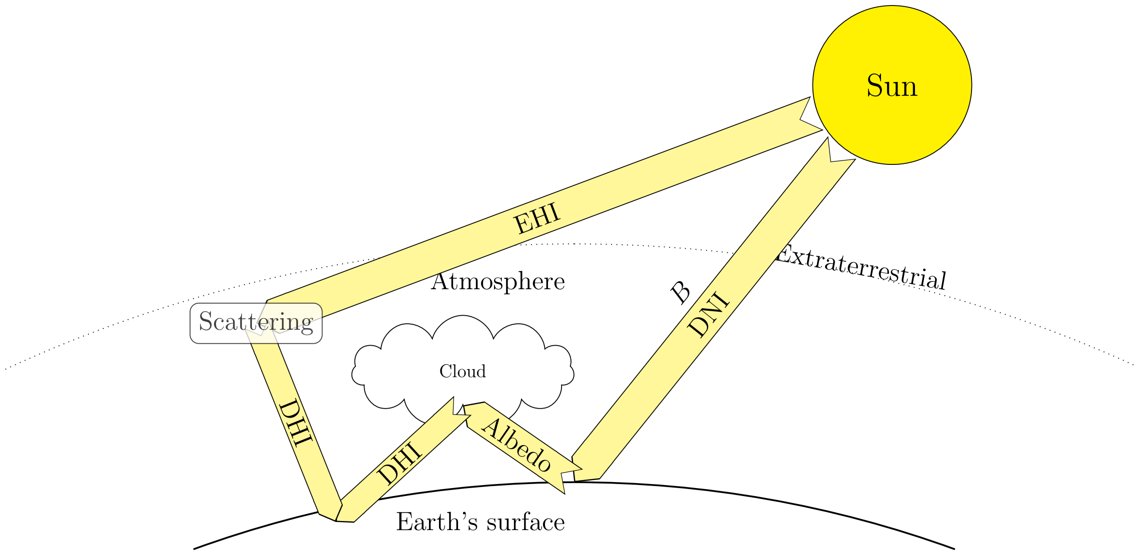

Figure 2.1: Energy from the sun takes multiple paths to the surface of the Earth.

Component	Summary	Symbol
EHI	Irradiance at top of the atmosphere	I_E
DNI	Irradiance directly from the sun	I_B
DHI	Irradiance from the sky	I_D
Albedo	Irradiance reflected by ground	-
GHI	Sum of DHI, DNI, and Albedo (usually negligible)	I_G
CSI	Maximum GHI measured during clear skies	I_C

Table 2.1

Summary of solar irradiance terminology. *Image Credit: Alex Hirzel*

2.2 Solar Forecasting Models

Solar radiation at the top of the atmosphere is constant over time, while the solar irradiance that reaches any point on earth's surface is a function of atmospheric and weather conditions above the location of interest. Cloud cover, aerosol and dust particles absorb and scatter radiation as solar irradiance passes through the atmosphere. All solar irradiance forecast models essentially offer a means to capture this relationship.

The section briefly reviews existing widely used models for estimating solar irradiance on the surface of the earth, with special consideration given to identifying the temporal and spatial domain of the forecast horizons (See Table 2.2).

Models	**Spatial**	**Temporal**
Persistence	Point Scale	Short Term
Sky Camera	Microscale & Mesoscale	Medium Term
Satellite Based	Mesoscale	Short & Medium Term
ARIMA	Microscale	Short & Medium term
Radiative	Microscale & Mesoscale	Medium & Long term
Empirical	Microscale & Mesoscale	Long Term
ANN	Microscale	Short & Medium term
NWP	Mesoscale & Global	Medium & Long term

Table 2.2
Spatial and Temporal domains of solar forecasting models. Adapted from [3, 7, 8]

The terms that define spatial and temporal domains for the rest of this report are described in Table 2.3 and Table 2.4.

Temporal Domain	**Forecast Range**
Short Term	0 h - 6 h
Medium Term	6 h - 24 h
Long Term	24 h - 72 h

Table 2.3
Summary of solar irradiance terminology

Spatial Domain	Forecast Range (radius)
Point Scale	Single site, usually < 0.01 km
MicroScale	0 km - 1 km
Mesoscale	1 km - 10 km
Global	> 10 km

Table 2.4
Summary of solar irradiance terminology

2.2.1 Persistence Models

Persistence models are naive forecast models predicated on the assumption that the solar irradiance at the current time step is likely to *persist* for the next time-step

$$I_{G_{t-1}} = I_{G_t}. \tag{2.1}$$

Their precision and accuracy decreases with forecast duration, and are known to be best suited for very short-term and near term (< 1 hour) forecasts [3] at a point-scale spatial domain. Persistence models are most useful for benchmarking other models. A more advanced model may not offer much value if it cannot out-perform a trivial persistence model.

2.2.2 Empirical Models

Empirical models are solar irradiance models based largely on empirical observations, and are not described through any mathematical or physical relationship between the inputs to the models.

2.2.2.1 Sunshine Based Models

First proposed by Ansgtrom [30] in 1924, sunshine duration based models establish a simple regressive relationship for the ratio of average daily GHI, I_G, to CSI, I_c, as a function of the ratio between the average daily sunshine duration, S_d, to the maximum sunshine duration, S_o, for a particular location

$$\frac{I_G}{I_c} = a_a + b_a\left(\frac{S_d}{S_o}\right), \tag{2.2}$$

where a and b are linear Angstrom model constants. Alternatively, Prescott [31] defined the linear relationship in terms of EHI, denoted as I_o

$$\frac{I_G}{I_o} = a_p + b_p\left(\frac{S_d}{S_o}\right). \tag{2.3}$$

While the regression constants a and b are empirically derived from measurements made from ground level stations, they have a physical significance. The variable a represents the overall atmospheric transmission on completely overcast days when $S_d/S_o = 0$. On a completely clear sky day, when $S_d/S_o = 1$, the sum of the terms $(a + b)$ is theoretically equal to 1. In the Prescott model, the sum $(a + b)$ represents the fraction of radiation received on clear sky days while accounting for dispersion of solar irradiation due to atmospheric effects.

The values of a and b, and additional emperical constants c and d for higher order equations have been developed for many locations across the world as summarized in Table 2.5. Detailed reviews of global solar radiation modeling using sunshine duration models for many more locations are available in [32] and [33].

In [34], a model for estimating the DNI was postulated as

$$\frac{I_b}{I_o} = cos(\alpha_s)\frac{S_d}{S_o}, \tag{2.4}$$

where α_s is the solar elevation angle, but hasn't found widespread application, aside from a few locations[35, 36, 37].

Quadratic [38] and higher order regressive [39] [40] models were developed to mitigate the sensitivity of linear models to periods of extreme cloud conditions, overcast $(S_d/S_o \approx 0)$ or clear-sky $(S_d/S_o \approx 1)$ [41].

There is no consensus on the benefits of second or third order regressive models over linear Angstrom models. Higher order regression relationships where shown to outperform linear models for some locations [42] [43] [44] [45] [46], performed the same as linear models in some locations [47], [48] [49] [50] and produced mixed results for yet other locations [51] [52]. A complete review of sunshine duration based quadratic regression models is available in[53].

Location	a	b	c	d
Algeria [54]	0.309	0.368	0	0
China [55]	0.2223	0.6529	0	0
Egypt [56]	0.228	0.527	0	0
India [47]	0.2281	0.5093	0	0
Italy [57]	0.117	0.692	0	0
Jordan [58]	0.174	0.615	0	0
Libya [45]	0.1000	0.8740	-0.255	0
Oman [46]	0.9428	-1.202	0.9336	0
Pakistan [59]	0.3480	0.3200	0.0700	0
Spain [60]	0.1840	0.6792	-0.113	0
Turkey[61]	0.2408	0.3625	0.4597	-0.3708
U.S.A[62]	0.81	-3.34	7.38	-4.51

Table 2.5
Values of constants in the higher order Angstrom-Prescott model empirically derived from measurements made at ground level stations for locations across the world

2.2.2.2 ASHRAE Models

The ASHRAE72 (American Society of Heating, Refrigerating and Air Conditioning Engineers, 1972) model [63] is a clear-sky model developed to estimate the monthly-average hourly GHI, I_G, incident on horizontal surface at sea level.

The model estimates DNI as a function of the zenith angle, z, and three time dependent parameters,

$$I_B = P \cdot R \cdot e^{-Q/\sec(z)}. \tag{2.5}$$

The values for constants P and Q were empirically derived from experimental data obtained from observation stations. ASHRAE provides values of all the constants for the 21st day of every month, along with a basic contour map of R (a non-parametric empirical constant) values for locations in the United States. The ASHRAE model can be extended to estimate GHI, and can also be adopted for locations at any altitude [64].

The updated ASHRAE2005 model[65] provided a calculation for parameterizing R using visibility index V - a variable measured at over 2000 ground weather monitoring weather stations [66] across the world.

In a comprehensive review of fifty-four clear-sky models conducted in 2012 [67], ASHRAE72 and ASHRAE2005 models were only two of the three models that only needed one input (zenith angle z) for computing hourly GHI and DHI values. Yet, ASHRAE models were only two of the fifteen models to meet the stringent criteria for a 'good model' (5% <Mean Bias Error (MBE) <+5% and RMSE<15%) across fifty-two rigorous stages of testing. In a similar comprehensive review of fifty-four models to compute the hourly DHI conducted in 2013 [68], ASHRAE2005 emerged as the best model and ASHRAE72 model emerged as the third best model to meet the stringent criteria for a 'good model' (10% <MBE <+10% and RMSE<30%) across eighteen rigorous stages of testing.

2.2.3 Temperature Based Models

Air temperature data is measured at practically every single meteorological station in the world. Inspired by relative success in incorporating temperature data in crop simulation models, [69] advocate for using using 'simple and robust' temperature measurements for solar irradiance forecasting

$$\frac{I_G}{I_o} = a + bt_{max} + ct_{max} + dN + e, \tag{2.6}$$

where N is the cloudiness Index, and a, b, c, d and e are constants whose values are available for most locations in the world.

Some authors [70] [71] [72] [73] have proposed location-agnostic analytics derivations for the constants that explicitly relate all the coefficients to climatological variables. While such models were shown to produce results comparable to other empirical models for a few locations [61], their performance at any arbitrary location is as yet remained un-examined.

Despite the limitations of location-agnostic models, empirical models for obtaining the monthly average daily GHI can be implemented for most locations in the world. The average daily sunshine duration, and air temperature are routinely measured at most meteorological stations across the world[66].

The prevalence of easily obtainable modeling and ease of computing GHI make empirical models the ideal first step in providing long term system energy output estimation and optimal solar energy system design and sizing.

2.2.4 Radiative Models

Even in ideal conditions with no cloud cover, a significant portion (25%) [74] of all EHI is lost before it reaches the ground. Irradiance on a clear day is well characterized as a function of elevation, ground angle, and the optical density of the atmosphere. Aerosol concentration, water vapor, turbulence, particulate matter and other atmospheric factors that contribute to adsorption or scattering of irradiance as it passes through the atmosphere [27, 75].

Clear-sky radiative models attempt to account for these effects using Radiative Transfer Modeling (RTM) - remote sensing instruments on satellites or ground measuring stations that model irradiance as a function of altitude, location and atmospheric conditions on clear sunny days.

2.2.4.1 The SOLIS and Ineichen Model

The launch of the Meteosat Second Generation (MSG) satellite in 2005 provided improved possibilities of monitoring Earth's atmosphere in real time [76]. MSG had higher spatial (km) and temporal (15 minute) resolution than previous satellites. MSG was also equipped with the ability to retrieve additional atmospheric parameters like cloud, water vapor, ozone and aerosol parameters.

The Solis clear-sky model describes a method of using the improved capabilities of the MSG satellite derived data along with RTM to obtain spectrally resolved global irradiance data. The hourly DNI was given by

$$I_B = I_o \cdot e^{\left(-m \cdot \tau\right)}, \tag{2.7}$$

where τ is the optical depth and m is the air mass.

While this relationship is valid for uniform path lengths of monochromatic radiation, the Solis model was extended to the multiple wavelength bands of incoming solar radiation [77]. The DNI, GHI and DHI are given respectively by

$$I_B = I_o \cdot exp\left(-\frac{\tau_b}{\sin^{\bar{b}} \alpha_s}\right) \tag{2.8a}$$

$$I_G = I_o \cdot exp\left(-\frac{\tau_g}{\sin^{\bar{g}} \alpha_s}\right) \cdot \sin \alpha_s \tag{2.8b}$$

$$I_D = I_o \cdot exp\left(-\frac{\tau_d}{\sin^{\bar{d}} \alpha_s}\right) \tag{2.8c}$$

where τ_b, τ_g and τ_d are beam, global, and diffuse total optical depths, $\bar{b}$, $\bar{g}$ and $\bar{d}$ are the corresponding fitting parameters obtained from RTM α_s is the solar elevation angle.

Instead of using monthly averaged parametric constants for atmospheric conditions used in the earlier ASHRAE and sunshine duration based models, the Solis model obtains high resolution diurnal clear-sky irradiance values. This is made possible by integrating radiative transfer model (RTM) calculations on the atmospheric parameters obtained at 100 km × 100 km resolution every day with the cloud information obtained through satellites at a 15 km × 15 km resolution every fifteen minutes.

The Solis model was shown [77] to be flexible to changes in the atmospheric state, and hence provides solar irradiance data with higher accuracy, temporal and spatial resolution in real time than earlier models.

The Ineichen model [78] was proposed as an analytic approximation of the SOLIS clear-sky model. Owing to the time consuming nature of obtaining beam, global and diffuse optical depths (τ_b, τ_g and τ_d) and the corresponding fitting parameters ($\bar{b}$, $\bar{g}$ and $\bar{d}$) in real time, the Ineichen Model provides analytical derivations of the fitting parameters. The inputs to the model are aerosol optical depth and water vapor column, both of which can be obtained for more than 500 locations in the world through the Aerosol Robotic Network (AERONET) database [79].

A complete derivation for the parametrization of different coefficients is available in [78]. The DNI, GHI and DHI values obtained from the simplified Ineichen model are shown[78] to compare well (RMSE of 1% for GHI, 2% for DNI and 5% for DHI) with the original Solis Model.

In a comprehensive review of eighteen radiative models for solar resource mapping performed in 2012 [80], the Ineichen model emerged as the second best model to forecast DNI within the uncertainty of broadband measurements from the best instruments with state-of-the-art calibration and radiometric techniques. The Ineichen model was also shown to have 'universal validity' due to its accurate and reproducible results though it only uses three inputs whereas other models require 5 or more atmospheric inputs.

2.2.5 Time Series Models

A time series is a series of observations collected at regular successive intervals. Time series models attempt to forecast trends in a data variable based on statistical modeling of observed patterns in the past [81]. Solar irradiation has distinct daily and seasonal time-varying components.

2.2.5.1 ARMA Models

Auto Regressive Moving Average (ARMA) models are a class of mathematical models with applications in fields with a large amount of historical data like finance, engineering and statistics. ARMA models provide a statistical description of stationary stochastic processes, using polynomials of two parts, an auto-regressive (AR) part and a moving average (MA) part (see Equation 2.9).

$$I_t = \underbrace{\sum_{i=1}^{p} \phi_i I_{t-i}}_{AR} + \underbrace{\sum_{j=1}^{q} \theta_i \epsilon_{t-i}}_{MA}, \tag{2.9}$$

where I can be any component of solar irradiance (I_G,I_D,I_E,I_B or I_C), ϕ and θ are the parameters of the model, and ϵ is the white noise error term. Equation 2.9 describes a time dependent variable X at any time-interval, as a combination of weighted sum of p most recent AR terms, and sum of q most recent random variations from the average over q terms. Equation 2.9 can be employed to forecast solar irradiance terms using procedures described in [82, 83].

1. Identification - Specify response series, run stationarity tests and identify candidate ARMA models
2. Estimation - Estimate the order a, b and parameters ϕ, ϵ, a non-trivial problem and an active area of research.
3. Forecasting - After model is trained, specify forecast intervals

2.2.5.2 ARIMA Models

ARMA models assume that time series models are stationary, i.e., the joint conditional probability distribution of the stochastic process is assumed to remain constant even when shifted in time. In reality, many real world time series data, including solar irradiance, have a time-dependent trends and periodicity. Solar irradiance varies through the day, and has seasonal variability.

A method to forecast non-stationary time series was first proposed by [84] in the form of Auto Regressive Integrated Moving Average (ARIMA) models. ARIMA models are a general class of statistical models that can be converted to stationary stochastic process using 'differencing'. ARIMA models are a generalization of ARMA models.

$$I_t^d = \underbrace{\sum_{i=1}^{p} \phi_i I_{t-i}^d}_{AR} + \underbrace{\sum_{j=1}^{q} \theta_i \epsilon_{t-i}}_{MA}, \tag{2.10}$$

where I^d is the component of solar irradiance, rendered stationary by differentiating d times, ϕ and θ are the parameters of the model, and ϵ is the white noise error term. The term I^d is the integrated back to recover the forecast for I.

ARIMA models can be adopted to forecast solar irradiance using a procedure described in [85].

1. Plot the data, identify patterns, and unusual observations.
2. Transform data to stabilize variance
3. Difference until irradiance data is stationary
4. Select appropriate AR and MA models
5. Consider appropriate seasonal models appropriate for seasonality of solar irradiance through the year
6. Calculate forecasts

2.2.6 Artificial Neural Network Models

Artificial Neural Networks (ANN) are general purpose computational intelligence machines that can be trained to learn, and subsequently recognize patterns in data. Computational ANNS are inspired by biological neurons. In ways similar to how neurons function, ANNs are collection of small interconnected processing units where information is passed along their interconnections [86]. The network acquired knowledge through a learning knowledge using the weights on the interconnections to store information.

The topography of a generic multilayer feedforward ANN is shown in Fig 2.2. Each neuron is connected to the neurons of the previous layer through 'adaptive sunaptic weights' [86]. In the learning phase, sets of inputs and outputs are passed to the network. The networks attempts to match the input data with the output data, adjusting weights to match the desired output. If the output does not match within a specific tolerance, the training algorithm adjusts the weights to reduce error. Over multiple passes of training data, the system learns to identify meaningful patterns in the input-output pairing. The most popular training algorithms used in ANNs is the back-prorogation algorithm, where the algorithm tries reduce total error by modifying node interconnection weights along a gradient.

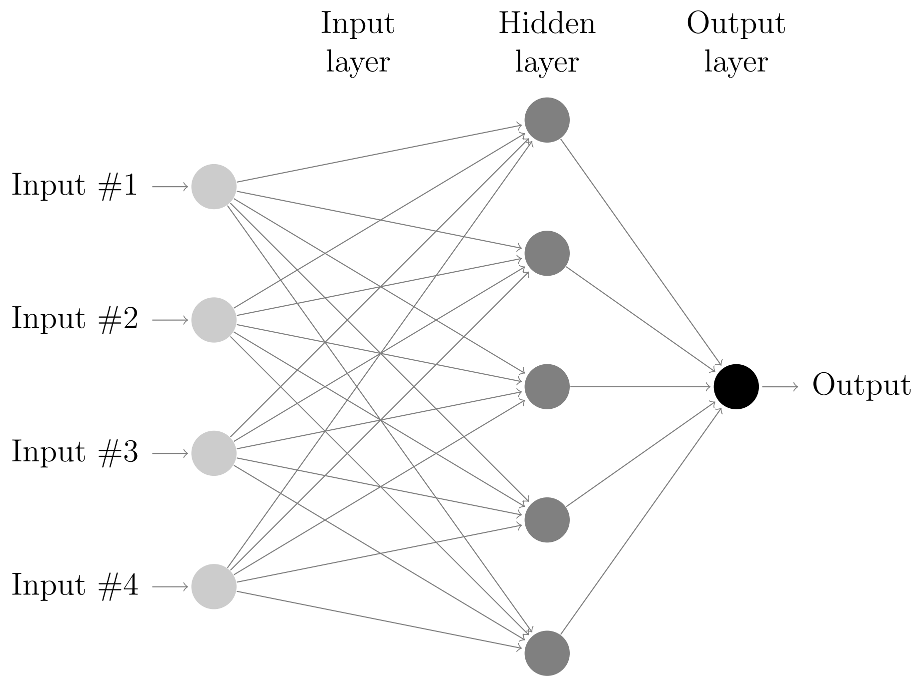

Figure 2.2: Model of an ANN, (*image modified from original code by Kjell Magne Fauske*)

ANN have been successfully applied in many mathematical and statistical domains including regression analysis, data mining and classification, processing, robotics, numeric control and game theory. Real world applications include medical diagnosis, image recognition, text mining, and more recently solar irradiation forecasting.

Reviews of solar irradiance forecasting widely available in literature [87, 88, 89, 90] describe most relevant input parameters for accurate solar forecasting. Models have used combinations of day of the year, air temperature, relative humdity, latitude, longitude, altitude, rainfall, cloud cover, wind-speed and atmospheric pressure. In addition, models also use archived irradiance data in combination with other archived data to perform time series forecasts. Models also range in diversity of model topography, with different input layers, network size and numbers of nodes at each layer.

ANN models, by their very nature, are 'black-box' models. It is difficult to interpret the causal or statistical relationship between single or multiple input parameters to the output. ANN models biased to the quality and quantity of training data available, and are prone to overfitting. A model trained for a specific location with a particular network configuration might not work for a different location or network parameters.

Despite these limitations, ANN models offer a number of advantages over traditional methods of statistical regression. ANN can model complex non-linear relationships between input and output variables, detect patterns in data without the need for parametric models or variables. ANNs are also very easy to implement with most modern computational software.

Recent advances in machine learning approaches, like feature selection and transfer learning [91] can help resolve the tradeoff between models that can generalize and over-fit to a set of input and model parameters.

2.2.7 Cloud Imagery Models

Cloud imagery models are predicated on the fact solar irradiance on the surface of earth (GHI) has an inverse relationship with the amount of cloud cover [92]. The amount of cloud cover can be estimated through satellites or ground-based measurement stations.

2.2.7.1 Satellite Derived Models

Progress in the fields of atmospheric radiative transfer modeling (RTM) has permitted greater predictive accuracy [93] by using parameterization of individual atmospheric factors.

The REST2 (Reference Evaluation of Solar Transmittance, 2 bands) model [93] forecasts DNI without the need for the computational complexity of spectral radiative models. Much like the Ineichen model, the REST2 obtains a parameter for spectral transmittance due to aerosols.

The DNI is obtained from a product of individual transmittances as show in 2.11

$$I_B = T_R T_g T_o T_n T_w T_a I_o, \tag{2.11}$$

where T_R, T_g, T_o, T_n, T_w, T_a are the transmittances due to Rayleigh Scattering, uniformly mixed gases absorption, ozone absorption, nitrogen dioxide absorption, water vapor absorption, and aerosol extinction respectively.

Inputs to this model are derived from radiometric measurements made from a network of more than 500 sites across the world through the AERONET [79] database. Details of such calculations, along with calculations for GHI and DHI, are available in [93].

In a comprehensive review of 18 radiative models for solar resource mapping performed in 2012 [80], REST2 emerged as the best model to forecast DNI. A condensed version [94] of the REST2 model was developed for easier estimation of worldwide clear-sky radiation data. In [94], the condensed REST2 model was shown be in 'excellent agreement' with the full REST2 model.

Input parameters for the REST2 model are available for more than 5400 stations around the world. The wide availability of model input parameters makes REST2 a good candidate for short-term operational planning, and the parametrization of physically-observed atmospheric conditions without the need for time intensive computational processes making the condensed REST2 model a minimal input option for long-term planning scenarios.

2.2.7.2 Sky Imagers

Sky Imagers are ground based cloud monitoring stations that offer intra-hour and sub-kilometer solar irradiance forecasts. Sky imaging instrumentation constitutes a rapid image capture camera equipped with a fish-eye lens, enclosed in a protective environmental housing. Most sky-imagers come equipped with in built in image processing algorithms to detect and forecast cloud movement. The spatial resolution depends on the number of pixels on the imaging camera, the position of the sun in the sky, topography, cloud distance and height. The temporal resolution depends on the operational ability of the image processing algorithms [95].

Irradiance forecasting using sky imagers occurs through three steps -

1. **Cloud Decision Algorithm** - Sky imaging cameras take a series of images that are processed through cloud decision algorithms. Due to Rayleigh scattering phenomena, clear skies appear blue. Individual pixels of the raw images are compared against a clear sky library. Pixels with a red-blue ratio higher than the threshold in the library are classified as a cloud.

2. **Cloud Motion** - Cloud motion is determined by cross-correlation of two consecutive sky images. Some post-processing elimiates erroneously small motion vectors, and the remaining vectors are averaged to obtain an averaged velocity and direction of cloud motion.

3. **Cloud Forecasting** - Cloud cover at any time t_0 is estimated from the cloud position and average cloud motion at t_{-1} obtained in the previous two steps.

2.2.8 Numerical Weather Prediction Models

Numerical Weather Prediction (NWP) models use mathematical models and computer simulations to forecast weather variables several hours in advance for large tracts of earth's surface. NWP models treat the atmosphere as a fluid, using thermodynamics to estimate the state of the fluid in the near future. The current state of weather is used as inputs to the model that produces about 125 output variables including GHI.

Out of many available NWP models, the North American Model (NAM), Global Forecast System (GFS) and European Center for Medium-Range Weather Forecast (ECMWF) are most commonly used for solar irradiance forecasting.

NWP Models	**NAM**	**GFS**	**ECMWF**
Spatial Domain	Continental US	Global	Global
Spatial Resolution	0.11°	0.5°	0.25°
Temporal Resolution	1 h	3 h	6 h
Forecast Horizon	36 h	180 h	240 h
Forecasts per day	4	4	2
MBE	57.5	37.4	31.4
RMSE	134.2	110.5	123.2

Table 2.6
Summary of NWP forecast models, adapted from [9]

2.2.8.1 NAM

National Oceanic and Atmospheric Administrations (NOAA) publishes the NAM forecast for (12 km x 12 km) grids spanning the entire continental United State. NAM forecast is a composite of over 125 weather variables. GHI forecast is made available through an internal assimilation of radiative transfer models, affected only by the atmospheric conditions present directly above each grid point.

This forecast is published four times every day, providing hourly output available up to 36 hours ahead. In addition, forecast of a 3-hour temporal resolution is available up to 84 hours ahead.

In a comprehensive review [9], NAM forecasts were found to be less accurate than other NWP models, despite its specialized application to the continental United States. (See Table 2.6)

2.2.8.2 GFS

GFS is a global forecast, also published by NOAA. GFS has a larger spatial resolution, larger temporal resolution and a longer forecast horizon than NAM forecasts.

Similar to NAM, GHI forecasts in GFS are obtained through radiative transfer models that account for the attenuating effect of absorption of solar radiation as it passes through earth's atmosphere.

2.2.8.3 ECMWF

Similar to GFS, ECMWF is a global forecast published by an intergovernmental consortium based in Europe. ECMWF provides medium range forecasts at larger spatial resolution, larger temporal resolution and a longer forecast horizon than both GFS and NAM forecasts.

Despite this, in a comprehensive review [9], ECMWF forecasts were found to be more accurate than other NWP models (See Table 2.6)

Medium-term solar forecasts are generally regarded to be more accurate when derived from NWP models. However, NWP models face significant challenges in forecasting solar irradiance in cloudy conditions due to limitations posed by complex cloud microphysics and real-time computation of their radiative properties. Consequently, NWP models are expected to show inherent regional biases due to micro-climatic conditions [9].

The temporal and spatial domains of the forecasting models are summarized in Table 2.2 and Figure 2.3.

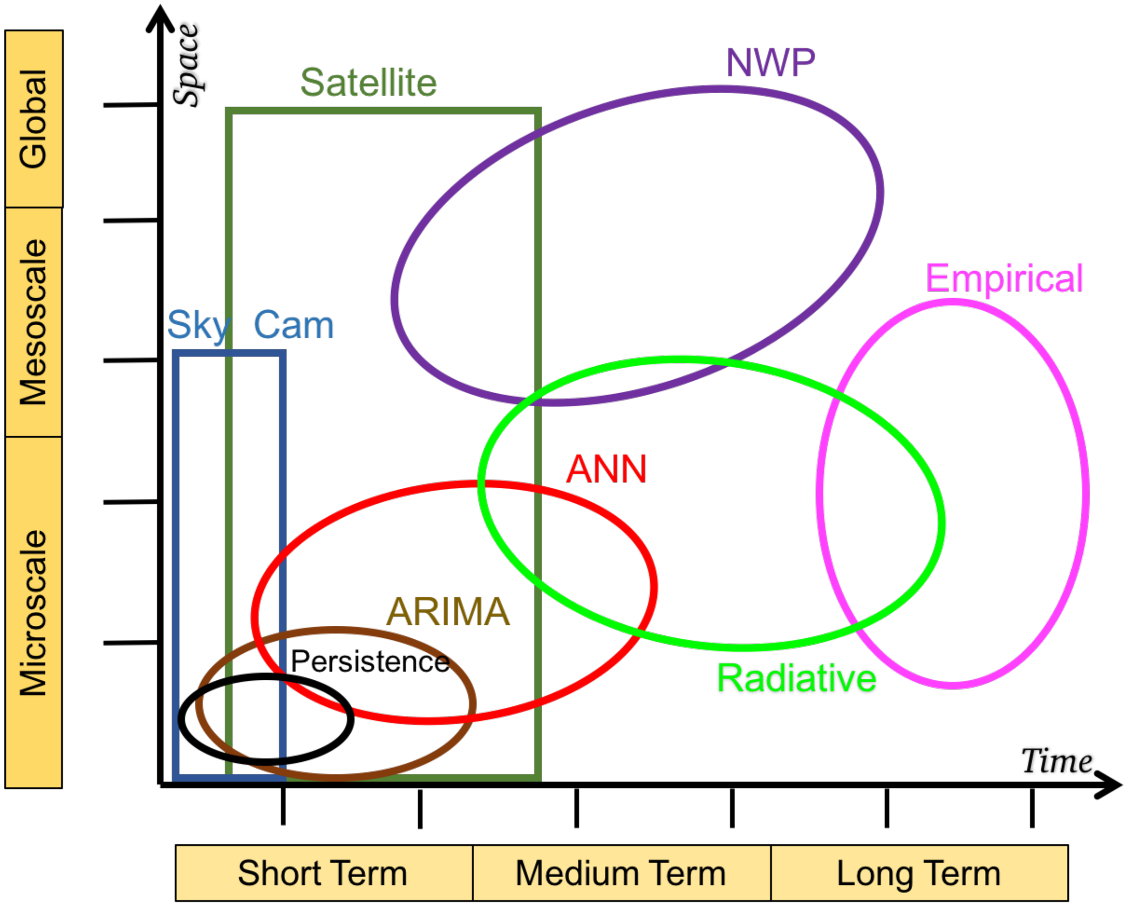

Figure 2.3: Spatial and temporal domains of solar irradiance models, adapted partly from [3]

2.3 Forecast Metrics

Accurate solar forecast facilitates more efficient integration of solar into utility resources, reduce curtailment of solar grid resources. Forecasts provide advance warning to render smooth transitions between solar and other energy resources.

To understand the value that solar forecasting provides, it is important to develop metrics to measure and assess the impact of forecasts have on integrating solar into the national grid [4].

The United States' Department of Energy (DOE) Sunshot Initiative identified [96] criteria to evaluate the usefulness of a forecasting metric:

- † simple and easily understandable
- † provides actionable insights
- † input data is manageable and acquirable, and
- † practical to make operational and planning decisions.

The Sunshot Initiative further identified a suite of comprehensive value-based and custom-developed metrics for Solar forecasting metrics for a diverse set of applications, forecast horizons and spatial domains. These metrics can be broadly divided into the following categories -

- † **Statistical** metrics are useful for evaluating the overall performance of forecasts, and for quick comparison between the performance of different forecast models.

Type	**Metric**	**Description**
Statistical	Pearson's coefficient	Linear correlation between forecast and measured solar irradiance
	RMSE	Evaluating overall accuracy, penalizing large errors (square)
	RMQE	Evaluating overall accuracy of the forecasts, penalizing large forecast errors (quadratic)
	MaxAE	Evaluating largest forecast error
	MAPE	Evaluating uniform forecast error
	KSI	Evaluating statistical similarity between forecast and measured
	MBE	Evaluating forecast bias
	OVERPer	Evaluating statistical similarity between large forecast errors
	Skewness	Evaluating assymetry of forecast error distribution
	Kurtosis	Magnitude of peak forecast distribution
Uncertainity metrics	Renyi entropy Standard Deviation	Quantifies uncertainity of forecast
Ramp Characterization	Swinging door algorithm	Extracts ramps in solar power output
Economic Metrics	95th percentile of forecast errors	Amount of non-spinning reserves to compensate for forecast errors

Table 2.7

Solar forecasting metrics adopted from [4, 5, 6]

† **Uncertainty Quantification** metrics measure the amount of deviation, variation or uncertainty of a forecast

† **Ramp Characteristics** identify the start and end points of ramps in solar power output

† **Economic** metrics measure the costs of maintaining spinning-reserves to compensate for inaccurate forecast errors.

A complete listing of applicable error metrics is described in Table 2.7

2.4 Applications & End-Users

Accurate solar forecasting has applications offering value to multiple stakeholders in the electric grid. Long-term forecasts of utility-scale solar may be used for reliability planning and scheduling of generation sources. Medium-term forecasts of roof-top solar at the distribution end may be employed in forecasting demand within a load-serving entity's service territory. Competitive electric markets may use short-term solar forecasts for bidding and trading energy services. Table 2.8 offers a comprehensive review of solar forecasting applications, end users and temporal domains.

Users	**Applications**	**Time-Horizon**
ISO/RTO	Reliability planning	Long Term
	Congestion management	Medium Term
	Unit commitment & dispatch	Short & Medium Term
	Load-flow, ramps & curtailment	Short & Medium Term
	Security. maintenance & outage	Medium & Long Term
Distribution Utilities	System planning	Long Term
	Outage management	Medium Term
	Load forecasting	Medium Term
	Smart Grid management	Short & Medium Term
Load Serving Entity	Scheduling & balancing	Short & Medium Term
Energy Traders	Bidding strategies	Short & Medium Term
Research labs Project developers	Integration & simulation studies	All Terms

Table 2.8
Adapted from [10, 11, 12]

Independent Systems Operators (ISO) and Regional Transmission Organizations) are independent neutral parties that responsible for the management and control of electricity across multiple states in the US territory, under regulation by the Federal Energy Regulation Commission (FERC). RTOs typically perform the same tasks as ISOs, but over a larger service territory. ISO/RTO administer regional wholesale electricity markets and plan for long term system reliability. Load serving entities are typically electricity utility companies that have vertically integrated generation, transmission and distribution services, while distribution utilities are smaller utilities that lack generation capabilities, either due to local energy policy or lack of adequate capacity. Energy traders are utilities or third party agents that participate in wholesale energy markets by buying and selling energy services.

The ultimate end-users may vary based on application, region and the nature of the regulation of electricity markets.

2.5 Summary

Models for accurate forecasts of solar irradiance integrates knowledge from diverse disciplines like atmospheric science, cloud physics, statistical mechanics, remote sensing, machine learning, and data mining. This chapter reviewed Various models for forecasting solar irradiance have been developed, each with its own model inputs, spatial and temporal domains.

In addition, this chapter also briefly reviewed grey literature discussing metrics for comparing and evaluating solar forecasts, applications and end users of accurate solar forecasts.

As the body of research in solar forecasting grows, there is a need to develop a shared and common understanding of this domain. The next chapter reviews the use of ontologies to facilitate knowledge sharing and reuse of information by agents interacting in the domain of solar forecasting.

Chapter 3

Ontology and Ontology Development Methodology

The term 'ontology' comes from Greek works *ontos for 'being' and* logos for 'word'. Ontology as a philosophical discipline has long been defined as a 'systematic explanation of being'. Led by Plato and followed by his student Aristotle, it is know as a study of objects, properties, events, processes and relations in all aspects of being [97].

3.1 Ontologies

Modern human activity is interdisciplinary, with people, organizations and software agents operating in a common space. Due to differing technical backgrounds, expertise, knowledge hierarchies, agents in the same environment may lack a shared understanding of the domain in which they interact [98]. As an example, solar forecasting models integrate knowledge from fields like cloud physics, statistical mechanics, artificial intelligence, machine learning and statistics. Software experts developing applications for the smart grid may lack an understanding of the underlying concepts and terminologies used in solar forecasting models. End users like energy market bidders and load serving utilities may have their own approach to structuring and organizing information and data that might not be congruent with the input requirements of solar forecasting models. Project developers may have goals, needs and expectations from solar forecasts that existing forecast models may not be able to meet.

Such differences in the definitions of concepts, structures, objects and relationships leads to poor communication, limiting inter-operability and reduces the potential for sharing knowledge and information about the domain of interest.

In recent years, ontologies have emerged as means to formally model the knowledge and structure of a system. Ontologies represent knowledge of a domain as a set of concepts, and relationships between pairs of concepts [20]. Ontologies are also a formal conceptualization that expresses a shared view of a domain between different parties [19] or a way of describing hierarchies or taxonomies in classified knowledge networks, expressed in a formal machine readable format.

As expressed in [20], the general goals of developing an ontology are to

† share a common understanding of the structure of information among human and software agents,

† enable reuse of information and domain knowledge,

† make domain assumptions and descriptions underlying an implementation explicit,

† separate domain knowledge from operational knowledge, and

† formally analyze terms and relationships that constitute domain knowledge to foster reuse and extension.

In addition, ontologies facilitate communication among humans and software agents without semantic ambiguity and provide foundations for inter-operability. Formal ontology development saves time and effort in building similar knowledge systems, and clarify the structure of knowledge through explicit domain knowledge definitions [99].

3.2 Ontology Language

The features of an ontology can also vary by the language used to describe the ontology. This thesis uses OWL, a standard ontology language from the World Wide Web Consortium (W3C). OWL is a formal semantic language that offers a rich set of differential logic operators (e.g., intersection, union) that allows complex concepts and relationships to be built from simpler concepts.

In OWL, any domain can be modeled through a shared vocabulary of individuals, properties and classes.

3.2.1 Individuals

Individuals represent objects in the domain of interest. In OWL, individuals have to be explicitly defined as same or different to each other. In Fig 3.1, Artificial Neural Network model (ANN), Global Horizontal Irradiance (GHI) and Cloud Cover are all instances or objects in the knowledge domain of solar forecasting.

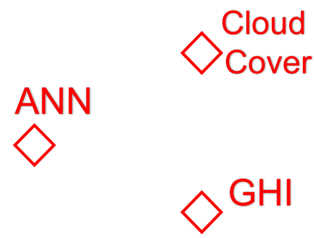

Figure 3.1: Representation of individuals in the solar forecasting ontology

3.2.2 Properties

Properties describe features and attributes of individuals. OWL offers the following type of properties

- † Datatype properties - describe relationships between objects and data values
- † Annotation properties - used for adding metadata, version information
- † Object properties - binary relationships that link two individuals. In OWL, object properties may have the following characteristics
  - – Functional - at most one individual is related to another individual through the property. In Fig 3.2, *hasOutput* is a property that describes the relationship between ANN and GHI, i.e., ANN *hasOutput* GHI.
  - – Inverse functional - the inverse of a property is functional. For example, in Fig 3.2 the property *isInputOf* is the functional inverse of the property *hasInput*.

– Transitive - if a transitive property relates individual a to b, and the same property relates b to c, then the a and c are related through the same property.

– Symmetric - if a symmetric property relates a to b, then the same property relates b to a.

– Asymmetric - if a property relates a to b, then b cannot be related to a the same property.

– Reflexive - when a property relates an individual to itself.

– Irreflexive - when a property cannot relate an individual to itself.

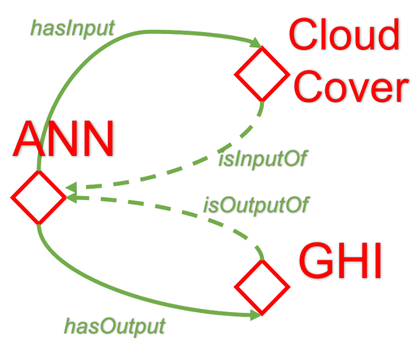

Figure 3.2: Example of properties that establish relationships between individuals in the solar forecasting ontology

3.2.3 Classes

Classes are the basic building blocks of an OWL ontology. Classes are sets that contain individuals whose descriptions precisely describe the class membership requirements.

Classes can have sub-classes that represent specification in addition to the what they may inherit from their super-class.

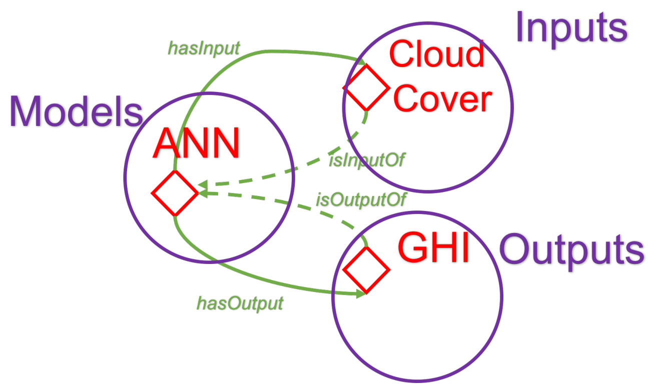

Figure 3.3: Examples of individuals, properties and classes in the OWL solar forecasting ontology

OWL lets classes be defined by the relationships between individuals. In Fig 3.3 ANN is a member of a class 'Models', where 'Models' is a set of individuals that are connected to another individual through the property *hasInputs*. OWL also supports the creation of anonymous classes through a restriction on object properties. These concepts are explored in a greater detail in Chapter 4

3.3 Ontology Development Methodologies

Many approaches have been proposed for formally developing ontologies. The rest of this section reviews common and widely used ontology development methodologies.

3.3.1 Uschold and King

The methodology by Uschold and King is among the first published ontology development methods. The approach is proposed in a set of stages -

1. **Identification** - Identify domain vocabulary, purpose, intended use, end-users, scope, terms. Define a set of competency questions.

2. **Building**

   † Ontology capture - Identify and textually define concepts and relationships.

   † Ontology coding - Transform words into formal ontology language using an ontology editor

   † Integrate existing ontologies

3. **Evaluation** Verify if classes, attributes and instances of classes meet requirements and purpose specified during the first stage.

4. **Documentation** - Document all the results from previous stages to aid in next iteration of ontology development.

3.3.2 SENSUS

SENSUS is a natural language based ontology developed to provide a broad conceptual structure for work in machine translation [24]. In contrast to other early ontology development methodologies which were developed from scratch, SENSUS was developed by extracting information from existing electronic resources.

Rather than develop a step by step or iterative process, SENSUS outlines the following principles for practitioners to bear in mind while designing an ontology -

† Do not over-commit on representational choices

† Should be extensible

† Should be extended based on needs identified during actual use

† Should integrate horizontally with other ontologies

† Structure ontology on organizing principles, conceptual clustering

3.3.3 METHODONTOLGY

METHODONTOLOGY uniquely identifies the process of knowledge acquisition, documentation and evaluation as being activities within an iterative 'life-cycle' of ontology development.

The life-cycle is described in the following stages -

1. **Specification** - Identify domain vocabulary, purpose, intended use, end-users, scope, terms, granularity, etc.
2. **Conceptualization** - Structure domain knowledge as a conceptual model using domain vocabulary. Build Glossary of Terms, concepts, verbs, class attributes and instances of classes.
3. **Integration** - Inspect meta-ontologies and reuse existing ontologies
4. **Implementation** - Use a development environment that looks for the most appropriate definitions, detects incompleteness, inconsistencies and redundant knowledge.

The following activities occur concurrently with the iterative life-cycle process

† **Knowledge Acquisition** - brainstorming, interviews with experts, literature

review

- **Evaluation** - verification for correctness and validation to check if ontology adequately represents the system it was built for.

- **Documentation** - Document as you go along in every stage.

3.3.4 On-To-Knowledge

On-To-Knowledge [100] introduces an ontology development methodology focused on bridging the gap between semantic descriptions of concepts for IT applications and human agents. This process is described in the following steps-

1. **Feasibility Study** - Identify problem and opportunity areas, potential solutions and put it in wider organizational perspective. Decision support for feasibility.

2. **Kickoff** - Generate requirement specification document containing

   - Goal

   - Domain

   - Applications supported

   - Knowledge sources

    - † Potential users and use-case scenarios
    - † Competency questions-concepts and relations
    - † Reusable ontologies

3. **Refinement** - Produce mature and application-oriented target ontology based on kick off specification
    - † Gather 'baseline taxonomy' or relevant concepts derived in kick off phase
    - † Develop 'seed ontology' after eliciting knowledge from domain experts
    - † Transfer seed to 'target ontology' using formal expression languages

4. **Evaluation** - Prove the usefulness of developed ontology
    - † Check if target ontology supports competency questions
    - † Test in the target application environment using feedback from beta users
    - † Trace usage patterns, iterate to identify areas most used or not used

5. **Maintenance** Change specifications based on developments in the real world. Cyclic refinement and evaluation phases

3.3.5 ONTOLOGY 101

Ontology 101 [20] presents an intuitive method for building ontologies. Unlike other methodologies, Ontology 101 is not presented as a step by step process, but instead as a set of stages that need to be accounted for through the complete iterative process.

- † Formulate competency questions to determine the domain and scope of the ontology.
- † Reuse existing ontologies
- † Build a Glossary of Terms
- † Identify classes from Glossary of Terms
- † Iterate over classes to identify properties and attributes
- † Specify range of values for attributes
- † Identify instances, run test case and repeat

As an informal methodology, Ontology 101 provides guidelines individuals with limited prior knowledge of onotology development. Due to its ease of implementation and widespread use, the rest of this report will use Ontology 101 methodology to develop solar forecasting ontology.

3.4 Summary

As described by [20], most ontologies follow come common principles ontology development is an iterative process, with no single and correct way to model a domain. The merits of an ontology rely on the the ability of the ontology developer to model ontological concepts semantically closely to objects and relationships in the domain of interest.

Many terms in the natural language lack a well defined semantic, particularly in interdisciplinary fields. Different people with different backgrounds and expertise may have different associations with the terms used to describe the knowledge of a domain. The challenges of achieving inter-operability rely on the skill of the ontology developer. Ontologies also need to be evolve with the domain, however regular updates and maintenance of the ontology can be a resource intensive.

Despite these limitations, explicit description of a domain of knowledge in a common vocabulary through an ontology remains a useful means for researchers to share information about a domain.

While there are advantages and challenges to each of the ontology development approaches, Ontology 101 is the only methodology described exclusively using Protégé software . In recent years, Protégé software has emerged as the leading ontology engineering tool, with over 300,000 users. Prot*ég*é software is a free, open source ontology editor that supports OWL semantic language (among others).

As an informal methodology, Ontology 101 provides guidelines individuals with limited prior knowledge of onotology development. Due to its relative ease of implementation and widespread use, the rest of this report will develop the solar forecasting ontology using Ontology 101 methodology. Due to its easy-to-use graphic user interface. the report will use Prot*ég*é software ,

Chapter 4

Ontology for Solar Forecasting

This chapter develops a solar forecasting ontology using the development methodology 'Ontology 101' [20]. Sections in this chapter may combine multiple steps in describing the development process.

4.1 Specification

The first step of developing an ontology is defining the domain and scope of the ontology. The solar forecasting ontology developed in this report integrates knowledge from academic papers, grey literature and expert reports on solar forecasting. The domain will model solar forecasting models, forecasting metrics, solar forecasting applications, end-users and the relationships between those concepts. This ontology is intended to be used by project developers to compare solar forecasting models, and choose most appropriate model for their use-case scenario, while working within the limitations of data and instrumentation availability.

The domain and scope of this ontology can be further defined by enumerating a list of competency questions that the ontology is expected to answer. A complete ontology should contain enough information to answer these questions, with the level of detail and representation necessary for the intended users of the ontology.

4.1.1 Competency Questions

While by no means exhaustive, Table 4.1 enumerates a sample of competency questions for illustrative purposes. In addition to identifying competency questions, Table 4.1 also picks out the important terms that need to be explained to the user. Our ontology will make 'statements' about these terms through relationships between them. At this stage, terms and relationships are expressed in the natural language without worrying about overlap of concepts.

Competency questions	**Class**	**Property**	**Class**
What do I need for the model to work?	ForecastModels	*hasInputs*	ForecastInputs
What models can I use this specific data for?	ForecastInputs	*isInputTo*	ForecastModels
What does my model tell me?	ForecastModels	*hasOutputs*	ForecastOutputs
How good is the forecast?	ForecastModels	*hasForecastMetrics*	ForecastMetrics
Can I use this model at this location?	ForecastModels	*hasGeographicAttribute*	Lat & Long.
Who uses solar forecasts? What uses?	EndUsers	*isResponsibleFor*	Applications

Table 4.1
A sample of competency questions

4.2 Related Ontologies

Reusing existing and validated ontology saves time and effort. Many concepts defined in an different ontology can be directly imported, with little modification, and applied to the domain of interest. Reusing existing ontologies may be an important requirement to ensure inter-operability among interdisciplinary and diverse ultimate users and their respective knowledge systems.

Many ontologies are designed for reuse under different contexts, with many publicly available ontologies specifically intended for use by the broader community. This section briefly reviews existing and publicly available ontologies that were considered for importing into the solar forecasting ontology.

4.2.1 Date and time

OWL - Time is a list of temporal concepts built for describing time related content of

Web pages. The ontology provides a vocabulary for expressing facts about topological relations among *instants* and *intervals* described using datetime information [101].

Our ontology imports some salient features of the OWL-Time ontology - Interval and Instant. Intervals are spans of time that have an beginning and an end, while Instants are intervals with zero length. Intervals can be used to express notions of time with regards to concepts of time like duration, overlaps, begins, ends, and finishes by.

Duration can be intervals of days, hours, seconds in terms of a predicate that enforces data property to specify time in units of seconds, minutes, hours, days, weeks, months and years. An interval can use multiple descriptions to describe the same duration in time.

Our ontology can be extended to include other OWL-Time concepts like TimeZone, which may be relevant for forecast models like NWP have to parse through forecast reports from a central repository.

4.2.2 Location

OWL - Basic-GEOis a Resource Description Foundation (RDF) vocabulary that provides semantic definitions for latitude, longitude and other spatial concepts. As a basic ontology, it does not cover the more complex concepts like polygon or boundary used by GIS services. While OWL- Basic-GEO was formally imported into our solar forecasting ontology, our representation for spatial concepts reuses the Basic-GEO formulation.

4.2.3 Units

Units of Measurement (OM) [102] is an OWL ontology of the domain of quantities and units of measure with particular reference to the fields of science and engineering. OM was designed to improve the annotation and interpretation of quantitative research data.

Each unit is expressed in terms of a set of base units depending on the SI system of units. Units are combined with prefixes such as milli or kilo to represent a multiplicative factor. For example, electromagnetic irradiance is defined in SI terms as $W\ M^{-2}$.

Our ontology imports features of OM most relevant for describing the domain of solar forecasting.

4.2.4 Weather

Many ontologies have been developed to describe concepts in the domain of weather and weather forecasting Weather Station, Weather-ONT and WeatherOntology. However, these ontologies were found to be too extensive for application in our solar forecasting ontology. Since these ontologies reused modified concepts of OWL-Time and Units, there was a risk of obfuscating domain knowledge due to overlapping classes, predicates and concepts.

4.2.5 Concentrated Solar Power

In [21], the authors describe a very simple ontology largely serves as a proof of concept for formally representing knowledge of solar radiation modeling and forecasting by the means of ontologies, with particular reference to concentrated solar power systems. While domain and scope of this ontology are somewhat limited, some concepts like temporal and spatial domain of forecast models were reused for creating our solar forecasting ontology.

4.3 Defining classes and hierarchy

Using competency questions motivated in 4.1, we enumerate a list of all terms and their properties we'd like to describe. Subsequently, we develop a class hierarchy and define the properties of the concepts. The features of an ontology can also vary by the language used to describe the ontology.

4.3.1 Class Hierarchy

From the list of terms generated in table 4.1, Prot*égé* software offers flexibility in describing the concepts either in terms of classes or individual instances. In this design methodology, terms that have independent existence are selected as classes.

Classes are organized into a hierarchical taxonomy, where subclasses inherit the properties of their superclass. An instance of a subclass by definition will be an instance of the superclass.

For example, the instance ANN in Fig 3.3 can also be expressed as a subclass ANN that inherits the properties and relationships of the class ForecastModels. Through inheritance, the class ANN also *hasInputs* some ForecastInputs and *hasOutputs* some ForecastOutputs. Here, instance ANN_1 may refer to a single particular ANN model

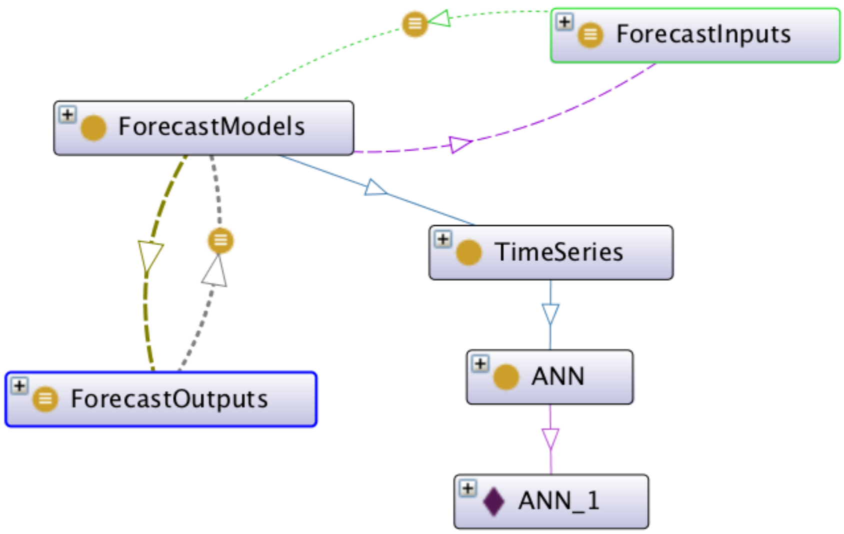

Figure 4.1: Refactoring some instances as classes and organizing them in a hierarchical taxonomy

as implemented in [2].

Terms	**Description**	**Attribute**
Forecast Models	Physical or statistical approximation of solar irradiance	Class
ANN	Artificial Neural Networks forecast models	Subclass to Class
ANN_1	ANN forecast model as implemented in [2]	Instance
Forecast Inputs	Inputs to a forecasting model	Class
Forecast Outputs	Outputs to a forecasting model	Class

Table 4.2
Glossary of terms resolved into classes and instances

4.4 Defining properties and relationships

OWL lets classes be defined by the relationships between concepts. In Table 2.2, short, medium and long terms are the temporal domains of forecast models. In

OWL, these concepts can represented by classes and relations as show in Table 4.3.

Class	**Relation**	**Class**
Persistence	*hasTemporalDomain*	Short Term
Sky Camera	*hasTemporalDomain*	Medium Term
Satellite Based	*hasTemporalDomain*	Short OR Medium Term
ARIMA	*hasTemporalDomain*	Short OR Medium Term
Radiative	*hasTemporalDomain*	Medium OR Long Term
Empirical	*hasTemporalDomain*	Long Term
ANN	*hasTemporalDomain*	Short OR Medium Term
NWP	*hasTemporalDomain*	Medium OR Long Term

Table 4.3
Temporal domains of solar forecasting models in Table 2.2 expressed as classes and relations. Adapted from [3, 7, 8]

These relationships are explicitly encoded in the development process of the ontology as shown in Fig 4.2, where models like NWP are subclasses of the class ForecastModels, but are connected to the class MediumTerm through the relation hasTemporalDomain.

4.5 Using Reasoners

In the hierarchical model of classification, the Prot*é*g*é* software environment that supports two kinds of hierarchies -

† Asserted hierarchy - Manually named and explicitly constructed.

† Inferred hierarchy - Infer facts that are not explicitly stated in the data model

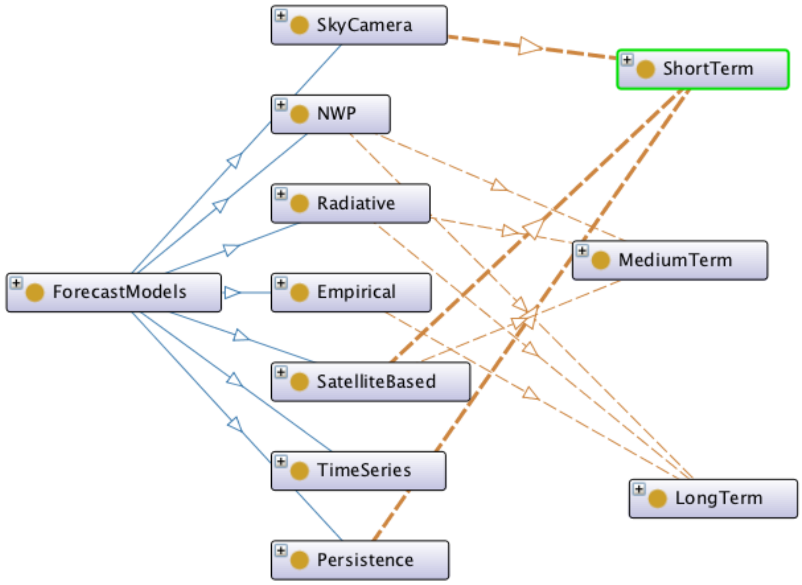

Figure 4.2: Defining relationships between classes

using reasoners.

Reasoners are a key component of working with OWL ontologies that use logical reasoning to test for consistency in asserted relations. Reasoners can test the membership of a class, classification hierarchy and check the logical consistence of an ontology. Advanced reasoners can also support automatic reasoning to generate inferred hierarchies.

In Prot*égé* software , consider the creation of a new subclasses of ForecastModels called ShortTermForecastModels. We *assert* that this subclass has a superclass called ForecastModels. The *Asserted* hierarchy of classes simply shows (see Fig 4.3 that ShortTermForecastModels is a subclass of ForecastModels, along with other classes

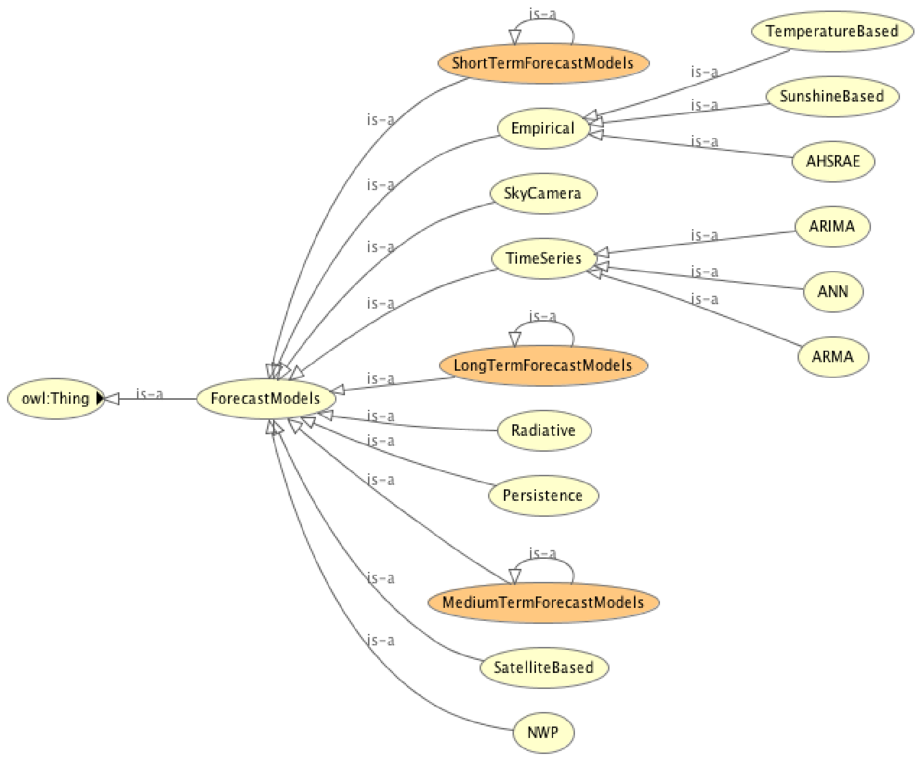

Figure 4.3: *Asserted* hierarchy of classes

explcitly defined earlier. We create two additional classes representing the temporal domains Medium Term and Long Term.

In addition, we specify the property that ShortTermForecastModels is equivalent to the class of ForecastModels that are connected to the class ShortTerm through the relation haveTemporalDomain. The default Prot*é*g*é* software Reasoner FACT++ supports automatic reasoning to generate inferred hierarchies as seen in Fig. 4.4.

Fig. 4.4 helps answer competency questions like - 'What models can I use for short

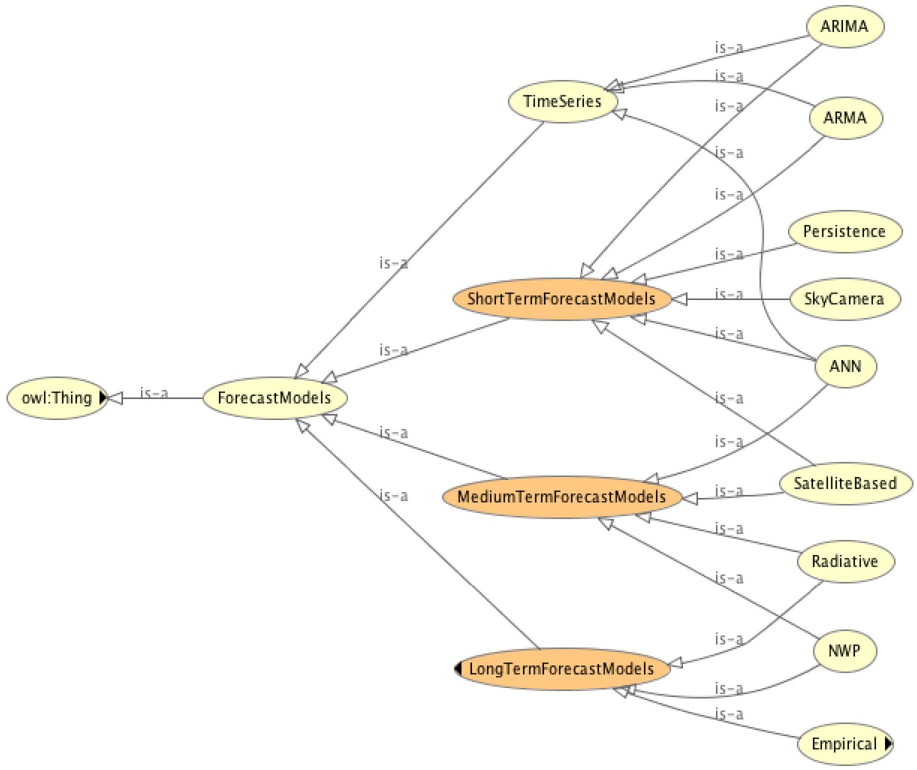

Figure 4.4: *Inferred* hierarchy of classes

term forecasting?' without explicitly encoding every single relationship in the knowledge model.

Similarly, the temporal domain relationships in Table ̃reftab:Applications can be encoded into the knowledge model. For demonstration purposes, we create a new dummy class called ForecastHorizon and a subclass called LongTermForecastHorizon. Here we specify that membership of this class is defined by the relation End-users, Applications or Forecast Models have temporal domain as Long Term. Since end-users do not have a temporal domain by default, they can inherit a relationship to the temporal domain by proxy through the application they are responsible for. The actual data property assertion in Protégé software is shown in Fig 4.5

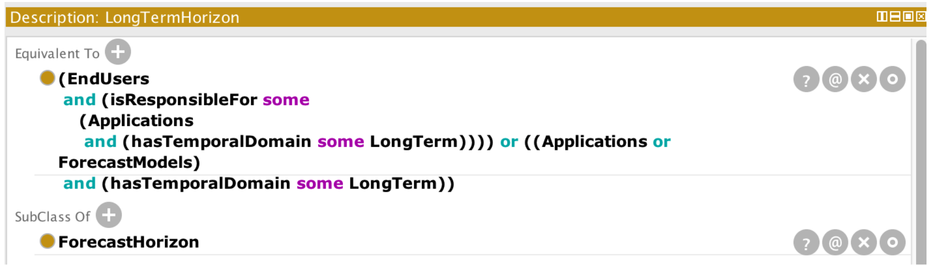

Figure 4.5: Class hierarchy relationship to identify dummy class LongTermForecastHorizon

The inferred hierarchy of this class is shown in Fig 4.6.

Fig 4.6 shows the relationships between all concepts in the domain based on their temporal horizon. Concepts can thus be members of named or anonymous subclasses based on their relationships inferred through the use of reasoners.

4.6 Domain Knowledge Validation by Use Case

Semantics of an ontology are verified through the use of reasoners to avoid overlap of concepts and relationships. Reasoners test for logical consistency, and if objects and properties are linked correctly based on defined rules and axioms.

Beyond such logical checks using reasoners, a systematic evaluation of an ontology can help users make informed decisions about choosing an ontology that best fits

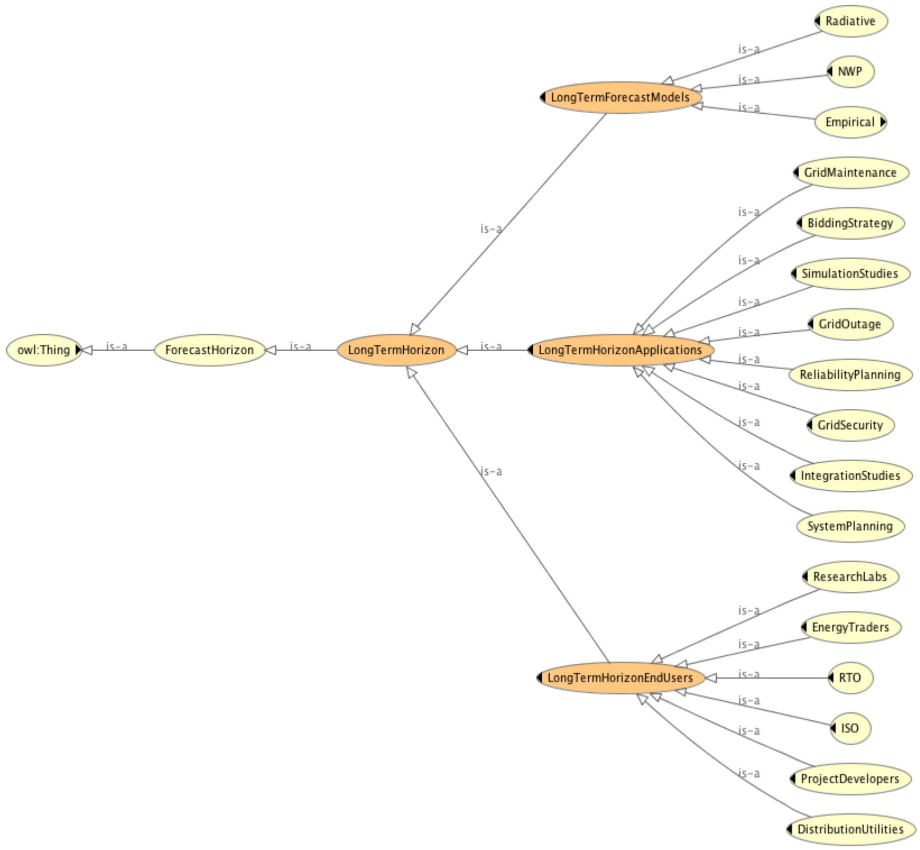

Figure 4.6: *Inferred* hierarchy of the dummy class LongTermForecastHorizon

their needs. Ontologies have to be further validated to test that they address the requirements that motivated their creation.

Validation of ontologies through illustrating use-cases is a common practice to determine if an ontology is accurate, adaptable and clear [103]. Accurate ontologies comply to the knowledge experts of the domain, and correctly represent the concepts of the world. Ontologies should be understandable, and offer a conceptual foundation for a range of anticipated uses.

The use-cases that follow are by no means exhaustive, but serve as a means to illustrate the quality of the ontology.

4.6.1 Identifying appropriate end-users based on constraint on forecast models

Consider a real world scenario where a software developer or research lab develops a solar forecasting model using Artificial Neural Networks. The researchers are interested in identifying the stakeholders in the smart-grid that would benefit from such a model.

We illustrate this use-case through a named class created for demonstration purposes. In the knowledge model, end-users are not explicitly related to forecast models. However, forecast models have temporal domains in which they are most effective. ANN are outperformed by other forecast models models for long term forecasts, and are most appropriate for short and medium term forecasts. In the ontology, ANN are connected to temporal domain through the *hasTemporalDomain* relationship.

Similarly, grid applications are connected to temporal domain through the *hasTemporalDomain* relationship. In Table 2.8, applications are within the domain of specific end users. In the ontology, end-users are connected to temporal domain through the *isResponsibleFor* relationship. For example, Independent Systems Operators (ISO) are directly responsible for unit commitment in the grid.

Therefore, the end users most likely to use ANN models are a subset, or subclass of all end users. For demonstration purposes, in the ontology a new named subclass called ANNEndUsers, a subclass under EndUsers, to represent this class of users. The users of this class were specified using the class hierarchy relationships shown in Fig 4.7

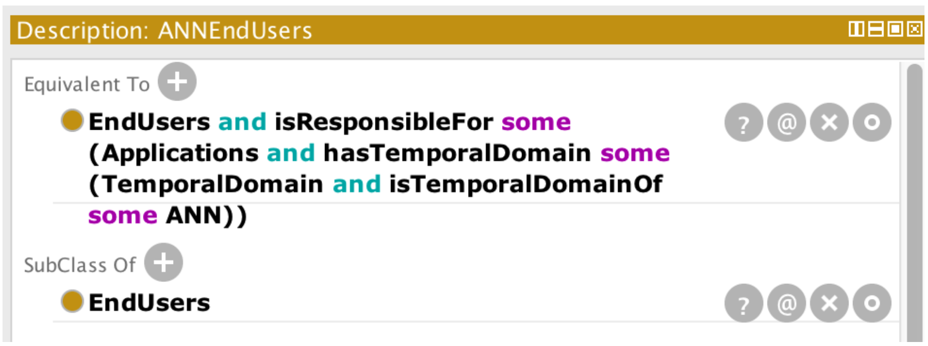

Figure 4.7: Class hierarchy relationship to identify dummy class of end users that may use ANN forecast models

The inferred hierarchy of this class is shown in Fig 4.8. ANNEndUsers is a subclass of all EndUsers who are likely to use ANN models.

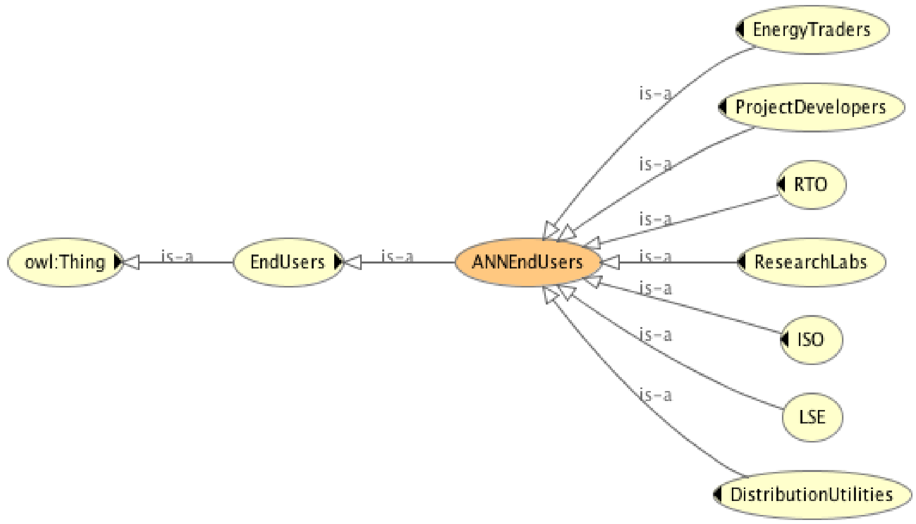

Figure 4.8: *Inferred* hierarchy of the dummy class ANNEndUsers, identifying the end users most likely to use ANN models.

4.6.2 Identifying appropriate applications based on constraint on available data

Consider a real world scenario where a project developer has no access to weather, atmospheric or archived power output from a solar generation site of interest. In this use case scenario, the developer is interested in identifying the grid applications that could benefit from forecasting solar irradiance using just parametric constants available for most locations in the US.

We illustrate this use-case through a named class created for demonstration purposes. In the knowledge model, applications are not explicitly related to input data for forecast models. However, in the ontology, input data is related to forecast models through the *isInputTo* relationship. Inversely, forecast models are connected to data through the inverse relationship *hasInput*. For example, AtmosphericData is connected to ANN models through *isInputTo* relationship.

As discussed earlier, forecasting models are connected to temporal domain through the *hasTemporalDomain* relationship, and grid applications are connected to temporal domain through the *hasTemporalDomain* relationship.

Therefore, the applications most benefit from using irradiance data to forecast solar irradiance, are subclass of all applications. For demonstration purposes, in the ontology a new named subclass called ParametricConstantsApplications, a subclass under Applications to represent this class of users. The members of this class were specified using the class hierarchy relationships shown in Fig 4.9

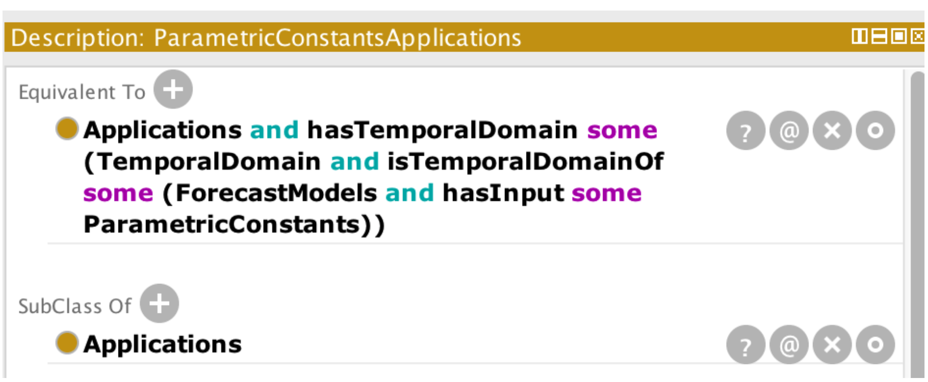

Figure 4.9: Class hierarchy relationship to identify dummy class of applications that may use solar irradiance forecast through parametric constants

The inferred hierarchy of this class is shown in Fig 4.10. ParametricConstantsApplications is a subclass of all EndUsers who are likely to use ANN models.

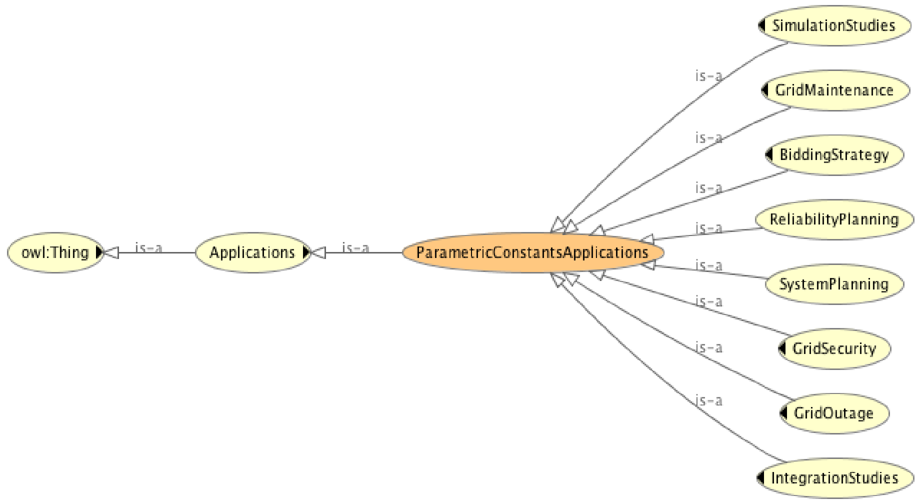

Figure 4.10: *Inferred* hierarchy of the dummy class ParametricConstantsApplications, identifying grid applications that can be addressed if only parametric constants were available as inputs to a class of solar irradiance models

4.6.3 Selecting appropriate models based on constraints on end-users

Consider a real world scenario where a software developer is developing solar forecasting tools for a specific end user, in this instance, a Load Serving Entity. In this use case scenario, the developer is interested in identifying all solar forecasting models that could be used to forecast solar irradiance according to parameters that could most benefit their client. At this stage of the development process, the developer does not have knowledge of specific grid applications. Subsequent to implementing forecasting models, the developer would like to identify the 'best' forecast models for his clients, where the criteria for comparing and evaluating merits of models is based on expert opinion, and widely accepted standards.

We illustrate this use-case through a named class created for demonstration purposes. In the knowledge model, forecast models are not explicitly related to end-users. However, in the ontology, In the ontology, forecast models are connected to temporal domain through the *hasTemporalDomain* relationship. Conversely, temporal domains are connected to forecast models through the inverse relationship *isTemporalDomainOf*.

Similarly, grid applications are connected to temporal domain through the *hasTemporalDomain* relationship. In Table 2.8, applications are within the domain of specific end users. In the ontology, end-users are connected to temporal domain through the *isResponsibleFor* relationship. Inversely, applications are connected to end-users through the inverse relationship *IsResponsibilityOf*. For example, unit commitment in the grid is the direct responsibility of Independent Systems Operators (ISO).

Therefore, the appropriate forecast models for the specific end user LSE are subclass of all end users. For demonstration purposes, in the ontology, a new named subclass called ModelsForLSE, a subclass under ForecastModels to represent this class of forecast models. The members of this class were specified using the class hierarchy relationships shown in Fig 4.11

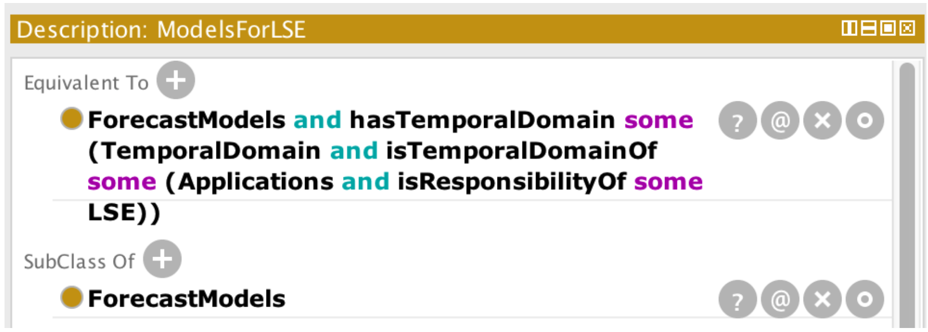

Figure 4.11: Class hierarchy relationship to identify dummy class of applications that may use solar irradiance forecast through parametric constants

The inferred hierarchy of this class is shown in Fig 4.10. ModelsForLSE is a subclass of all ForecastModels most appropriate for end users like LSEs.

Subsequent to identifying all forecasting models most appropriate for end users like LSEs, the developer can identify appropriate forecasting models based on available data similar to use-case 2 discussed earlier. When not constrained by choice of available data as inputs to forecast models, the developer can compare the outputs of forecasts from different models using widely accepted industry standards developed by experts at NREL [4, 5] and US DOE [6]. This knowledge is also modeled in the domain as an asserted class as show in Fig. 4.13.

A description for each of the error metrics is encoded in the ontology using annotations, and data property assertions are used to encode datatype values like percentage.

The class hierarchy asserted in Fig. 4.13 represents the current knowledge of experts

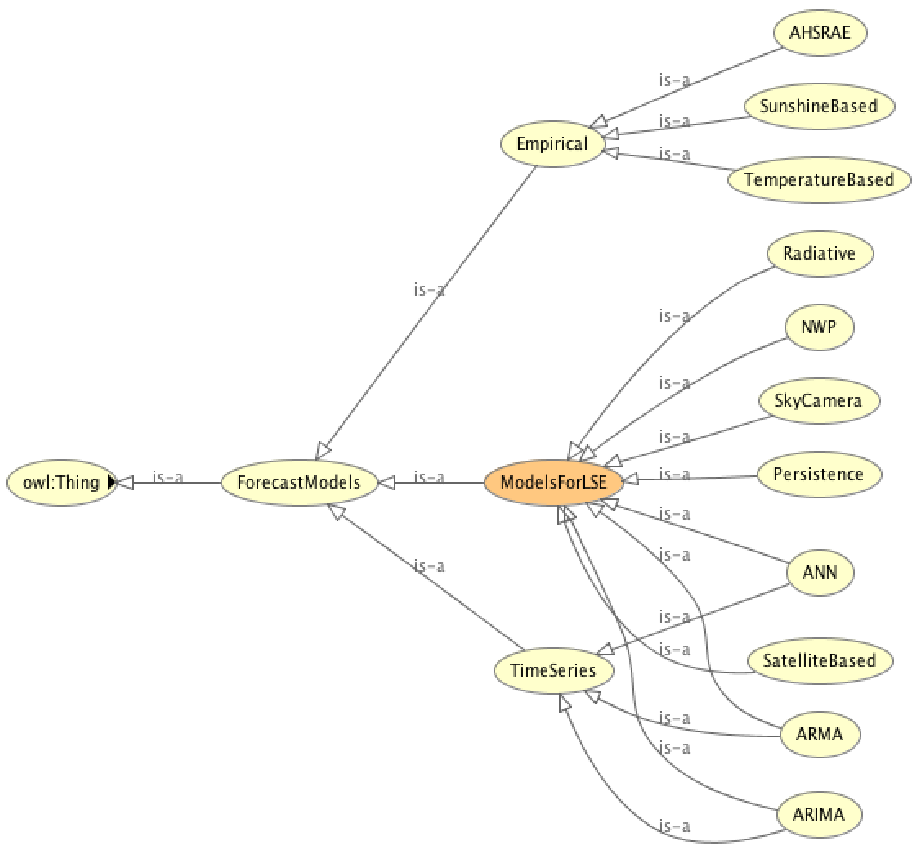

Figure 4.12: *Inferred* hierarchy of the dummy class ModelsForLSE is a subclass of all ForecastModels most appropriate for end users like LSEs.

in the domain. Metrics for evaluating solar forecasts is a currently evolving research topic, with the most recent suite of metrics published just a year before the writing of this report.

As knowledge of this field expands to identify most appropriate metrics for specific grid applications, our ontology can be easily extended to model that relationship using existing object property assertions and axioms.

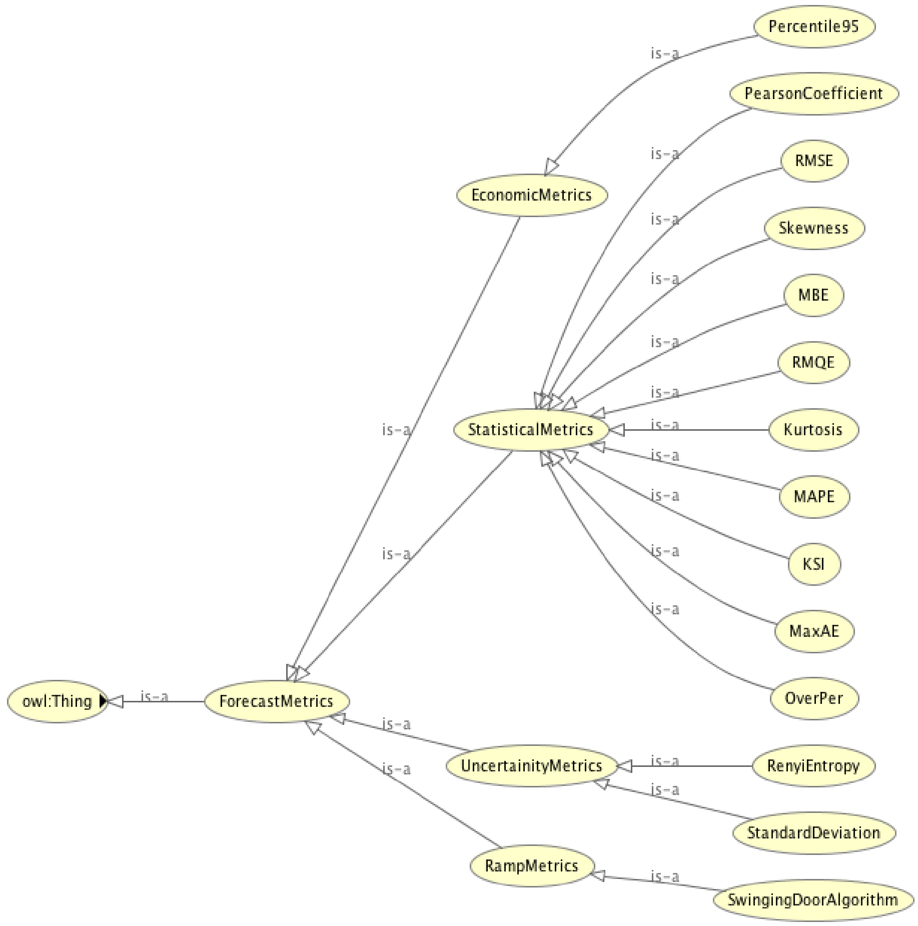

Figure 4.13: Asserted hierarchy of metrics for evaluating solar forecasts, adapted from NREL [4, 5] and US DOE [6]

4.6.4 Summary

This chapter develops SF-ONT, a formal ontology that maps the domain knowledge of solar irradiance forecasting using Ontology 101, an ontology development methodology. The ontology is expressed in OWL using the software package Prot*égé* software . The top-level concepts in solar irradiance forecasting are expressed in the form of a hierarchical taxonomy as shown in Figure 4.14. Relationships in classes are expressed through object properties, while data properties are used to express properties of individuals. Fig 4.15 is an overview of the comprehensive ontology.

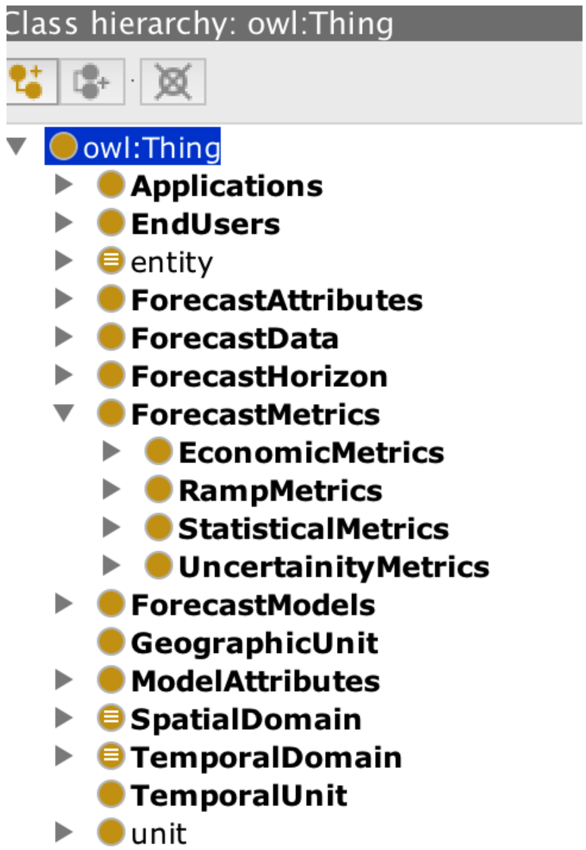

Figure 4.14: Top level concepts in solar irradiance forecasting expressed as classes in SF-ONT

SF-ONT is available for download and use at https://github.com/akantamn/sf-ont

The semantic and syntactic consistency of the ontology was tested using reasoners configured for the ontology development environment Prot$\acute{e}$g$\acute{e}$ software . Several use-cases were employed to illustrate how the ontology anticipates the requirements of the users. Use cases also validate the transfer from the real world to the semantic knowledge model of the ontology.

Ontology metrics:	
Metrics	
Axiom	2139
Logical axiom count	980
Declaration axioms count	439
Class count	160
Object property count	89
Data property count	11
Individual count	163
DL expressivity	SHOIF(D)
Class axioms	
SubClassOf	394
EquivalentClasses	40
DisjointClasses	5
GCI count	0
Hidden GCI Count	41

Figure 4.15: Summary of SF-ONT ontology metrics

The use cases described in this chapter are by no means exhaustive, and the ontology developed here is by no means 'complete'. Ontology development is by nature an iterative process.

At present, this ontology identifies all the top level concepts in the domain of solar irradiance forecasting, and models their relationships within the context of the current knowledge in the domain. As knowledge in the domain expands, this ontology can be extended from the basic building box, modified and maintained to suit the evolving needs of the users.

Chapter 5

Summary

Accurate forecasts of solar irradiance can help with the reliable and sustainable integration of solar energy resources into the national grid, providing value for many grid applications. In recent years, a growing body of academic research has developed models for forecasting solar irradiance, integrating knowledge from fields like atmospheric science, cloud physics, statistical mechanics, artificial intelligence and machine learning. Experts from national labs and industry have identified metrics for comparing solar forecasts, and described applications that will benefit from accurate solar forecasts.

Due to differing technical backgrounds, expertise, knowledge hierarchies, terminologies, technical knowledge, and expectations, the diverse stakeholders in the world of solar forecasting may lack a shared understanding of the domain in which they interact.

This report describes a step towards improving communication, inter-operability and the potential for sharing knowledge and information about solar forecasting, using ontologies.

Firstly, the report describes the basics of solar irradiance and forecasting, and reviews academic literature on recent advancements in improved solar forecasting models, with special reference to their spatial and temporal domains. The report also identifies grid applications and end-users that will benefit from accurate solar forecasts. The report also reviews recent developments in developing a suite of metrics for evaluating solar forecasts.

Subsequently, the report describes ontologies and briefly reviews methodologies for developing ontologies. Using Ontology 101, an ontology development methodology, the report then describes SF-ONT - a formal ontology that maps the knowledge domain of solar irradiance forecasting. SF-ONT is available for download and use at https://github.com/akantamn/sf-ont

Lastly, the report describes the testing and validating of solar forecasting ontology for accuracy and completeness using built-in *Protégé* software reasoners.

The ontology described in this report identifies all the top level concepts in the domain of solar irradiance forecasting, and models their relationships within the context of the current knowledge in the domain. As knowledge in the domain expands, this ontology can be extended from the basic building box, modified and maintained to suit the evolving needs of the users. As the ontology continues to be tested, validated and refined in practice by ultimate users, the shared vocabulary of the ontology can be useful for developing standards and practices. Even when standards are developed independently, ontologies can help interface between standards and practices from different expert fields relevant to the domain.

References

[1] SEIA. Solar energy facts: Q2 2015 Technical report, Solar Energy Industries Association, Washington DC, **2015**.

[2] Amrouche, B.; Le Pivert, X. *Applied energy* **2014**, *130*, 333–341.

[3] Diagne, M.; David, M.; Lauret, P.; Boland, J.; Schmutz, N. *Renewable and Sustainable Energy Reviews* **2013**, *27*, 65–76.

[4] Zhang, J.; Florita, A.; Hodge, B.-M.; Lu, S.; Hamann, H. F.; Banunarayanan, V.; Brockway, A. M. *Solar Energy* **2015**, *111*, 157–175.

[5] Zhang, J.; Hodge, B.-M.; Florita, A.; Lu, S.; Hamann, H. F.; Banunarayanan, V. In *3rd International workshop on integration of solar power into power systems, London, England*, 2013.

[6] Marquez, R.; Coimbra, C. F. *Journal of solar energy engineering* **2013**, *135*(1), 011016.

[7] Glassley, W.; Kleissl, J.; Shiu, H.; Huang, J.; Braun, G.; Holland, R. *California Institute for Energy and Environment* **2010**.

[8] Pelland, S.; Remund, J.; Kleissl, J.; Oozeki, T.; De Brabandere, K. *IEA PVPS, Task* **2013**, *14*.

[9] Mathiesen, P.; Kleissl, J. *Solar Energy* **2011**, *85*(5), 967–977.

[10] MAkhyoun, M. *Solar Electric Power Association* **2014**.

[11] Widiss, R.; Porter, K. *NREL Technical Report* **2014**, *303*, 275–3000.

[12] Zieher, M.; Lange, M.; Focken, M. *German Federal Ministry for Economic Cooperation and Development* **2015**.

[13] Reichelstein, S.; Yorston, M. *Energy Policy* **2013**, *55*, 117–127.

[14] Frankl, P.; Nowak, S. *Technology roadmap: solar photovoltaic energy;* OECD/IEA, 2014.

[15] Kostylev, V.; Pavlovski, A.; others. In *1st International Workshop on the Integration of Solar Power into Power Systems Aarhus, Denmark*, 2011.

[16] Letendre, S.; Makhyoun, M.; Taylor, M. *Solar Electric Power Association, Tech. Rep* **2014**.

[17] Inman, R. H.; Pedro, H. T.; Coimbra, C. F. *Progress in energy and combustion science* **2013**, *39*(6), 535–576.

[18] Gruber, T. R. *Knowledge acquisition* **1993**, *5*(2), 199–220.

[19] Guarino, N.; Oberle, D.; Staab, S. In *Handbook on ontologies;* Springer, 2009; pages 1–17.

[20] Ontology development 101: A guide to creating your first ontology. Noy, N. F.; McGuinness, D. L.; others. **2001**.

[21] Piazza, A.; Faso, G. In *Advances onto the Internet of Things;* Springer, 2014; pages 325–337.

[22] Zhou, Q.; Natarajan, S.; Simmhan, Y.; Prasanna, V. In *Information Technology: New Generations (ITNG), 2012 Ninth International Conference on*, pages 775–782. IEEE, 2012.

[23] Fernandez-Lopez, M.; Gomez-Perez, A.; Juristo, N. **1997**.

[24] Swartout, B.; Patil, R.; Knight, K.; Russ, T. In *Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, pages 138–148, 1996.

[25] Horridge, M.; Knublauch, H.; Rector, A.; Stevens, R.; Wroe, C. *University of Manchester* **2004**.

[26] McEvoy, A.; Markvart, T.; Castaner, L. *Practical Handbook of Photovoltaics, Second Edition: Fundamentals and Applications;* Academic Press, 2011.

[27] Glossary of Solar Radiation Resource Terms. Laboratory, N. R. E. **2013**.

[28] Paltridge, G. W.; Platt, C. M. R. *Radiative processes in meteorology and climatology / by G. W. Paltridge and C. M. R. Platt;* Elsevier Scientific Pub. Co Amsterdam ; New York, 1976.

[29] Estimating Monthly Solar Radiation in South-Central Chile. Álvarez, J.; Mitasova, H.; Allen, H. L. **2011**.

[30] Angstrom, A. *Quarterly Journal of the Royal Meteorological Society* **1924**, *50*(210), 121–126.

[31] Prescott, J. *Transactions of the Royal Society of South Australia* **1940**, *64*(1), 114–118.

[32] Besharat, F.; Dehghan, A. A.; Faghih, A. R. *Renewable and Sustainable Energy Reviews* **2013**, *21*, 798–821.

[33] Katiyar, A.; Pandey, C. *Journal of Renewable Energy* **2012**, *2013*.

[34] Suehrcke, H.; McCormick, P. *Solar Energy* **1989**, *42*(4), 303–309.

[35] Power, H. C. *Solar Energy* **2001**, *71*(4), 217–224.

[36] El-Metwally, M. *Atmospheric Research* **2004**, *69*(3), 217–239.

[37] Power, H. *Theoretical and Applied Climatology* **2003**, *76*(1-2), 47–63.

[38] Ögelman, H.; Ecevit, A.; Taşdemiroğlu, E. *Solar Energy* **1984**, *33*(6), 619–625.

[39] Samuel, T. *Solar Energy* **1991**, *47*(5), 333–337.

[40] Bahel, V.; Bakhsh, H.; Srinivasan, R. *Energy* **1987**, *12*(2), 131–135.

[41] Benson, R.; Paris, M.; Sherry, J.; Justus, C. *Solar Energy* **1984**, *32*(4), 523–535.

[42] Akinoğlu, B.; Ecevit, A. *Energy* **1990**, *15*(10), 865–872.

[43] Taşdemiroğlu, E.; Sever, R. *Energy Conversion and Management* **1991**, *31*(6), 599–600.

[44] Aksoy, B. *Renewable Energy* **1997**, *10*(4), 625–633.

[45] Said, R.; Mansor, M.; Abuain, T. *Renewable energy* **1998**, *14*(1), 221–227.

[46] Ampratwum, D. B.; Dorvlo, A. S. *Applied Energy* **1999**, *63*(3), 161–167.

[47] Katiyar, A.; Pandey, C. K. *Energy* **2010**, *35*(12), 5043–5048.

[48] Wu, Z.; Du, H.; Zhao, D.; Li, M.; Meng, X.; Zong, S. *Theoretical and Applied Climatology* **2012**, *108*(3-4), 495–503.

[49] Zhao, N.; Zeng, X.; Han, S. *Energy Conversion and Management* **2013**, *76*, 846–851.

[50] Rensheng, C.; Shihua, L.; Ersi, K.; Jianping, Y.; Xibin, J. *Energy Conversion and Management* **2006**, *47*(7), 865–878.

[51] Tarhan, S.; Sarı, A. *Energy Conversion and Management* **2005**, *46*(4), 605–613.

[52] Jin, Z.; Yezheng, W.; Gang, Y. *Energy Conversion and Management* **2005**, *46*(2), 257–268.

[53] Bakirci, K. *Renewable and Sustainable Energy Reviews* **2009**, *13*(9), 2580–2588.

[54] Chegaar, M.; Chibani, A. *Energy conversion and management* **2001**, *42*(8), 967–973.

[55] Li, H.; Ma, W.; Lian, Y.; Wang, X.; Zhao, L. *Renewable Energy* **2011**, *36*(11), 3141–3145.

[56] El-Metwally, M. *Journal of Atmospheric and Solar-Terrestrial Physics* **2005**, *67*(14), 1331–1342.

[57] Jain, P. *Solar & wind technology* **1986**, *3*(4), 323–328.

[58] Alsaad, M. *Solar & wind technology* **1990**, *7*(2), 261–266.

[59] Ahmad, F.; Ulfat, I. *Turkish Journal of Physics* **2004**, *28*(5).

[60] Almorox, J. y.; Hontoria, C. *Energy Conversion and Management* **2004**, *45*(9), 1529–1535.

[61] Ulgen, K.; Hepbasli, A. *International Journal of Energy Research* **2002**, *26*(5), 413–430.

[62] Lewis, G. *Energy conversion and management* **1992**, *33*(12), 1097–1099.

[63] Handbook, A. *American Society of Heating, Refrigerating and Air Conditioning Engineers, Atlanta* **1972**, pages 385–443.

[64] Wong, L.; Chow, W. *Applied Energy* **2001**, *69*(3), 191–224.

[65] Handbook, A. *Refrigerating and Air-Conditioning Engineers, Atlanta, Georgia* **2005**.

[66] Monthly climatic data for the world. Satelite, N.; Service, I.

[67] Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models. Badescu, V.; Gueymard, C. A.; Cheval, S.; Oprea, C.; Baciu, M.; Dumitrescu, A.; Iacobescu, F.; Milos, I.; Rada, C. **2012**.

[68] Badescu, V.; Gueymard, C. A.; Cheval, S.; Oprea, C.; Baciu, M.; Dumitrescu, A.; Iacobescu, F.; Milos, I.; Rada, C. *Theoretical and applied climatology* **2013**, *111*(3-4), 379–399.

[69] Paulescu, M. In *Modeling Solar Radiation at the Earths Surface;* Springer, 2008; pages 175–192.

[70] Glover, J.; McCulloch, J. *Quarterly Journal of the Royal Meteorological Society* **1958**, *84*(360), 172–175.

[71] Gariepy, J. *International Report, Service of Meteorology, Government of Quebec, Canada* **1980**.

[72] Rietveld, M. *Agricultural Meteorology* **1978**, *19*(2), 243–252.

[73] Njau, E. C. *Renewable energy* **1996**, *7*(1), 105–108.

[74] Frederick, J.; Hodge, A. *Atmospheric Chemistry and Physics* **2011**, *11*(3), 1177–1189.

[75] Reno, M.; Hansen, C.; Stein, J. *Sandia National Laboratories SAND2012-2389* **2012**.

[76] Schmetz, J.; Pili, P.; Tjemkes, S.; Just, D.; Kerkmann, J.; Rota, S.; Ratier, A. *Bulletin of the American Meteorological Society* **2002**, *83*(7), 977–992.

[77] Mueller, R.; Dagestad, K.-F.; Ineichen, P.; Schroedter-Homscheidt, M.; Cros, S.; Dumortier, D.; Kuhlemann, R.; Olseth, J.; Piernavieja, G.; Reise, C.; others. *Remote sensing of Environment* **2004**, *91*(2), 160–174.

[78] Ineichen, P. *Solar Energy* **2008**, *82*(8), 758–762.

[79] Holben, B.; Eck, T.; Slutsker, I.; Tanre, D.; Buis, J.; Setzer, A.; Vermote, E.; Reagan, J.; Kaufman, Y.; Nakajima, T.; others. *Remote sensing of environment* **1998**, *66*(1), 1–16.

[80] Gueymard, C. A. *Solar Energy* **2012**, *86*(8), 2145–2169.

[81] Prema, V.; Rao, K. U. *Renewable Energy* **2015**, *83*, 100–109.

[82] Rout, M.; Majhi, B.; Majhi, R.; Panda, G. *Journal of King Saud University-Computer and Information Sciences* **2014**, *26*(1), 7–18.

[83] Huang, R.; Huang, T.; Gadh, R.; Li, N. In *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, pages 528–533. IEEE, 2012.

[84] Box, G. E.; Jenkins, G. M.; Reinsel, G. C.; Ljung, G. M. *Time series analysis: forecasting and control;* John Wiley & Sons, 2015.

[85] Hyndman, R. J.; Athanasopoulos, G. *Forecasting: principles and practice;* OTexts, 2014.

[86] Kalogirou, S. A. *Renewable and sustainable energy reviews* **2001**, *5*(4), 373–401.

[87] Kalogirou, S. *Energy Conversion and Management* **1999**, *40*(10), 1073–1087.

[88] Mellit, A. *International Journal of Artificial intelligence and soft computing* **2008**, *1*(1), 52–76.

[89] Yadav, A. K.; Chandel, S. *Renewable and Sustainable Energy Reviews* **2014**, *33*, 772–781.

[90] Gardner, M. W.; Dorling, S. *Atmospheric environment* **1998**, *32*(14), 2627–2636.

[91] Pan, S. J.; Yang, Q. *IEEE Transactions on knowledge and data engineering* **2010**, *22*(10), 1345–1359.

[92] Kasten, F.; Czeplak, G. *Solar energy* **1980**, *24*(2), 177–189.

[93] Gueymard, C. A. *Solar Energy* **2008**, *82*(3), 272–285.

[94] Gueymard, C. A.; Thevenard, D. *Solar Energy* **2009**, *83*(11), 1998–2018.

[95] Chow, C. W.; Urquhart, B.; Lave, M.; Dominguez, A.; Kleissl, J.; Shields, J.; Washom, B. *Solar Energy* **2011**, *85*(11), 2881–2893.

[96] Solar forecasting metrics. U.S.DOE. Webinar, **2014**.

[97] Smith, B. In *Blackwell Guide to the Philosophy of Computing and Information*, pages 155–166. Oxford, 2003.

[98] Uschold, M.; Gruninger, M. *The knowledge engineering review* **1996**, *11*(02), 93–136.

[99] Fisseha, F. In *Nordic Agricultural Ontology Service (AOS) Workshop Royal Veterinary and Agricultural University Copenhagen, Denmark*, 2003.

[100] Sure, Y.; Staab, S.; Studer, R. In *Handbook on ontologies;* Springer, 2004; pages 117–132.

[101] Pan, F.; Hobbs, J. R. In *FLAIRS Conference*, Vol. 5, pages 560–565, 2005.

[102] Rijgersberg, H.; van Assem, M.; Top, J. *Semantic Web* **2013**, *4*(1), 3–13.

[103] Staab, S.; Studer, R. *Handbook on ontologies;* Springer Science & Business Media, 2013.